Construction of Invariant Measures of Lagrangian Maps I: Minimisation and Relaxation

Siniša Slijepčević

Abstract

If F is an exact symplectic map on the d-dimensional cylinder $\mathbb{T}^d \times \mathbb{R}^d$, with a generating function h having superlinear growth and uniform bounds on the second derivative, we construct a strictly gradient semiflow ϕ^* on the space of shift-invariant probability measures on the space of configurations $(\mathbb{R}^d)^{\mathbb{Z}}$. Stationary points of ϕ^* are invariant measures of F, and the rotation vector and all spectral invariants are invariants of ϕ^*. Using ϕ^* and the minimisation technique, we construct minimising measures with an arbitrary rotation vector $\rho \in \mathbb{R}^d$, and with an additional assumption that F is strongly monotone, we show that the support of every minimising measure is a graph of a Lipschitz function (the Birkhoff-Mather theorem for symplectic maps). Using ϕ^* and the relaxation technique, assuming a weak condition on ϕ^* (satisfied e.g. in Hedlund's counter-example) we show existence of double-recurrent orbits of F (and F-ergodic measures) with an arbitrary rotation vector $\rho \in \mathbb{R}^d$, and the action arbitrarily close to the minimal action $A(\rho)$.

1991 MSC: 58F05, 58F11

MSC 2000: 37J**, 37J50, 37L45

Siniša Slijepčević
Department of Mathematics
Bijenička 30
10000 Zagreb, Croatia

fax: +385 1 6601221
e-mail: `slijepce@cromath.math.hr`

1 Introduction

The duality between gradient and Hamiltonian dynamics has been a part of mathematical folklore for some time now. A beautiful example of an application of this idea is Floer's approach to Arnold's conjecture on the number of fixed points of a Hamiltonian map: he studied a "gradient" dynamics on the space of contractible paths, whose equilibria are the 1-periodic orbits of the Hamiltonian flow.

Here we construct a strictly gradient semiflow ϕ^* on a separable, metrisable, complete space $\mathcal{A}^*$ which in a way contains all the information about a given Hamiltonian (Lagrangian in particular) system. This means that, in principle, *a question about a Lagrangian system can be formulated as a question about its "dual" gradient semiflow* ϕ^*.

For the simplicity of arguments, here we restrict our attention to Lagrangian maps on the cotangent bundle $T^*\mathbb{T}^d$ of the d-dimensional torus $\mathbb{T}^d$. Lagrangian maps are the discrete-time analogues of Lagrangian flows, i.e. maps whose orbits are extremals of the action functional $\sum h(x_i, x_{i+1})$, where $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is the generating function. The semiflow ϕ^* is the semiflow on the space of shift-invariant measures on a subset of $(\mathbb{R}^d)^{\mathbb{Z}}$, induced by the formally gradient dynamics of the action functional. The set of equilibria of ϕ^* is the set of invariant measures of the Lagrangian map.

From the point of view of variational theory, the semiflow ϕ^* is a powerful new tool for construction of closed sets of measures, such that the minimum of the action functional restricted to the closed set is a local minimum of the action functional.

We develop two methods of use of the semiflow ϕ^* for construction of invariant measures of Lagrangian maps: the method of *minimisation* (a modification of the standard variational technique), and *relaxation*. We then apply these methods and address the following questions:

Minimising measures. The most successful generalisation of Aubry-Mather theory to more degrees of freedom is by Mather [Mat91], who constructed minimising measures of a Lagrangian flow with an arbitrary rotation vector, and proved that their supports are graphs of a Lipschitz function. Using the minimisation technique, we prove the analogues of Mather's results for Lagrangian maps. The use of the semiflow ϕ^* slightly simplifies and generalises the proof of the existence of action-minimising measures with a given rotation vector: positive definiteness of the generating function is not a necessary condition; periodicity, superlinear growth and e.g. uniform bounds on the second derivatives of h suffice.

Using the additional assumption that $x \mapsto h_2(x, y)$, $x \mapsto h_1(y, x)$ are diffeomorphisms (where indices denote partial derivatives), we prove the Birkhoff-Mather theorem: the support of a minimising measure is a graph of a Lipschitz function. It is a generalisation of the well-known Birkhoff theorem for twist maps; partial generalisation of Herman's result [Her89] who proved the same for measures supported on Lagrangian tori; and the analogue of Mather's theorem [Mat91] for positive-definite Lagrangian flows.

Quasiperiodic orbits. The Mather's construction does not generalise the main result of Aubry-Mather theory, and does not succeed to represent every rotation vector with an *ergodic* minimising measure (and with a minimising orbit of the Lagrangian flow with that rotation vector). Such an attempt, as was shown by Hedlund [Hed32], must fail. Levi [Lev97] succeeded to construct orbits (only locally minimising) with every rotation vector in Hedlund's counter-example. Here we generalise Levi's construction, and demonstrate that the method of relaxation is successful when constructing invariant measures which minimise the action functional only locally. We introduce the "distance" function d^* on the space of measures, with values in $[0, \infty]$, and prove the following theorem for Lagrangian maps:

Theorem *We denote the set of all ergodic measures of a given Lagrangian map F by $\mathcal{S}^*_E$, the set of all action-minimising invariant measures of F with the rotation vector ρ by $\mathcal{M}^*_\rho$, and by $A(\rho)$ the value of the action functional on $\mathcal{M}^*_\rho$. If there exists a neighbourhood $\mathcal{U}$ of $\mathcal{M}^*_\rho$ such that*

$d^*|_{\mathcal{U}} < \infty$, *then for every* $\epsilon > 0$, *there exists* $\mu \in \mathcal{S}^*_E$ *such that the action* $A(\mu) \leq A(\rho) + \epsilon$.

We then argue why we believe that the condition of the finiteness of the distance function in a small neighbourhood of $\mathcal{M}^*_\rho$ is typically (possibly always) satisfied.

The paper is structured as follows: In section 2 we discuss three standing assumptions on the generating function h of a given Lagrangian map: (A1) periodicity, (A2) the superlinear growth, and (A3) the existence of the semiflow ϕ^*. We show that (A3) is satisfied if e.g. the second derivatives of h are bounded, and in the case of twist maps. In section 3 we show that ϕ^* is strictly gradient, and that the "interesting" part of the phase space of ϕ^* is an increasing union of ϕ^*-invariant compact sets. The methods of minimisation and relaxation are developed in section 4. In section 5 we show that the invariants of the semiflow ϕ^* are: ergodicity, rotation vector, and all spectral invariants, including the Kolmogorov-Sinai entropy. Minimising measures are discussed in section 6, and quasi-periodic orbits in section 7. In section 8 we apply the results to the case of twist maps ($d = 1$), and discuss open problems.

2 Lagrangian maps

In the following, $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ will denote a C^2-function, the discrete "Lagrangian". We will impose three conditions on h which will be assumed throughout the paper. We denote by $|.|$ the Euclid metric on $\mathbb{R}^d$.

(A1) Periodicity. For each integer vector $a \in \mathbb{Z}^d$, and for each $x, y \in \mathbb{R}^d$, $h(x+a, y+a) = h(x, y)$.

(A2) Superlinear growth. $\lim_{|\eta|\to\infty} h(x, x+\eta)/|\eta| = +\infty$, uniformly in $|\eta|$.

The condition (A1) implies that, if h is a generating function of a map F, then F can be restricted to a map on a torus $\mathbb{T}^d$. The conditions (A1) and (A2) are the discrete-time analogues of the conditions on Lagrangian flows used by Mather in [Mat91].

2.1 Gradient dynamics of the action functional.

Before we state the third condition, we need to introduce the main technique of the paper. In the following, the elements of $(\mathbb{R}^d)^{\mathbb{Z}}$ are called configurations. We study the system of equations

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = h_2(u_{i-1}, u_i) + h_1(u_i, u_{i+1}), \quad i \in \mathbb{Z} \tag{1}$$

$$\mathbf{u}(0) = \mathbf{u}^0, \tag{2}$$

where $\mathbf{u}^0 \in (\mathbb{R}^d)^{\mathbb{Z}}$ is the initial condition, indices denote partial derivatives, and the solution is a function $\mathbf{u} : \mathbb{R}^+ \to (\mathbb{R}^d)^{\mathbb{Z}}$ satisfying (1), (2). We can formally write the equation (1) as

$$\mathbf{u}_t = -\nabla h(\mathbf{u}),$$

where $h(\mathbf{u}) = \sum_{i\in\mathbb{Z}} h(u_{i-1}, u_i)$; therefore we say that (1) is a formally gradient (or an extended gradient) system of equations.

Let $\tilde{\mathcal{B}}$ denote the set of all configurations $\mathbf{u} \in (\mathbb{R}^d)^{\mathbb{Z}}$, such that

$$\sup_{i\in\mathbb{Z}} |u_{i+1} - u_i| < \infty.$$

Proposition 2.1 *Assume that* h *is* C^2, *and that it satisfies (A1), (A2). Then the equation (1) generates a local flow on* $\tilde{\mathcal{B}}$ *(i.e. for each* $\mathbf{u}^0 \in \tilde{\mathcal{B}}$, *there exists* $\delta > 0$, *and a unique solution* $\mathbf{u} : (-\delta, \delta) \to \tilde{\mathcal{B}}$ *of (1), (2) for some* $\delta > 0$*).*

Proof: For each $\mathbf{u}^0 \in \tilde{\mathcal{B}}$, let the space $l_{\mathbf{u}^0,\infty}$ be the space of all $\mathbf{v} \in \mathbb{R}^{\mathbb{Z}}$ such that $\sup_{i\in\mathbb{Z}} |u_i - v_i| < \infty$, with the norm $||v||_{\mathbf{u}^0,\infty} = \sup_{i\in\mathbb{Z}} |u_i - v_i|$. Given $\mathbf{u}^0 \in \tilde{\mathcal{B}}$, (A1) implies that the right-hand of (1) is locally Lipschitz on $l_{\mathbf{u}^0,\infty}$. Well-known results on existence and uniqueness of solutions of ordinary differential equations on Banach spaces (see e.g. [DK74]) now imply the claim. $\square$

In other words, Proposition 2.1 implies existence of a local flow on $\tilde{\mathcal{B}}$. The third condition will be that the local flow can be extended to a semiflow on a suitable space which contains $\tilde{\mathcal{B}}$. Given a local flow or semiflow ϕ on a space $\mathcal{X}$, we use the notation $\phi^t(x) = \phi(t,x)$. We say that a semiflow ϕ is backward-unique, if for each t, $\phi^t : \mathcal{X} \to \mathcal{X}$ is injective. We say that a semiflow ϕ is continuous, if for each t, the function ϕ^t is continuous, and if for each $x \in \mathcal{X}$, the function $t \mapsto \phi(t,x)$, $t \in [0,\infty)$, is continuous.

We denote by T_a, $a \in \mathbb{Z}^d$, and S the translations defined on $(\mathbb{R}^d)$:

$$\begin{aligned} T_a(\mathbf{u})_i &= u_i + a, \\ S(\mathbf{u})_i &= u_{i-1}. \end{aligned}$$

The condition (A1), and the homogeneity in i of the equation (1) imply respectively that the solution of (1) commutes with the translations T_a, S. We denote by T the group of all translations T_a, $a \in \mathbb{Z}^d$.

In the following we always assume the topology on $\tilde{\mathcal{B}}$ induced by the product topology on $(\mathcal{R}^d)^{\mathbb{Z}}$, and $\mathcal{B} = \tilde{\mathcal{B}}/T$ is the quotient space with the quotient topology. For a given $C > 0$, $\tilde{\mathcal{B}}_C$ is the set of all $\mathbf{u} \in \mathbb{R}^{\mathbb{Z}}$ such that $\sup_{i\in\mathbb{Z}} |u_i - u_{i+1}| \leq C\}$.

We now assume the following:

(A3) Existence of a semiflow. There exists a complete separable metric space $\tilde{\mathcal{A}}$, such that for each $C > 0$, $\tilde{\mathcal{B}}_C$ is a subspace of $\tilde{\mathcal{A}}$, $\tilde{\mathcal{A}}$ a subset of $(\mathbb{R}^d)^{\mathbb{Z}}$, and such that:

(i) For each $i \in \mathbb{Z}$, the projection $\pi_i : \tilde{\mathcal{A}} \to \mathbb{R}$, $\phi_i(\mathbf{u}) = u_i$ is continuous;

(ii) For each $C > 0$, the set $\mathcal{Y}_C := \{\mathbf{u} \in \tilde{\mathcal{A}} \ : \ \forall k \in \mathbb{Z},\ |u_{k+1} - u_k| \leq C(k^2+1)\}/T$ is compact;

(iii) The translations S, T_a, $a \in \mathbb{Z}$, are homeomorphisms of $\tilde{\mathcal{A}}$;

(iv) the equation (1) generates a backward unique continuous semiflow $\tilde{\phi}$ on $\tilde{\mathcal{A}}$.

Remark 1: The conditions (A3),(i),(ii) and partially (iii) are conditions on the topology of $\tilde{\mathcal{A}}$, and say that this topology is "close" to the topology induced by the product topology on $(\mathbb{R}^d)^{\mathbb{Z}}$ in the following sense. Denote by τ the topology on $\tilde{\mathcal{A}}$. (A3),(i) is equivalent to the condition that the topology τ contains the induced product topology. If for each $C > 0$, the induced topology τ on $\mathcal{Y}_C$ is the same as the induced product topology on $\mathcal{Y}_C$, then the Tychonoff theorem and completeness of $\tilde{\mathcal{A}}$ imply (A3),(ii).

Standing assumption: In the rest of the paper, we assume that h is C^2, and that it satisfies (A1-3).

We set $\mathcal{A} = \tilde{\mathcal{A}}/T$, and denote by ϕ the semiflow on $\mathcal{A}$, whose lift on $\tilde{\mathcal{A}}$ is $\tilde{\phi}$. The condition (A1) implies that ϕ is well defined. In the following, we do not distinguish between elements of $\tilde{\mathcal{A}}$ and their equivalence classes in $\mathcal{A}$, and perform the calculations in $\mathcal{A}$ when possible.

The role of the condition (A3) is in a way parallel to the Mather's condition in [Mat91] of the completeness of the Euler-Lagrange flow, though the relationship of these two conditions remains unclear. We immediately give two simpler conditions which will imply (A3).

2.2 Examples.

Example 1. We can assume the following:

(E1) Assume that the second derivatives h_{11}, h_{12}, h_{22} are uniformly bounded on $\mathbb{R}^d \times \mathbb{R}^d$.

We choose any $\lambda > 1$, and define the Hilbert space $\mathcal{A}_\lambda$ of all configurations $\mathbf{u} \in (\mathbb{R}^d)^{\mathbb{Z}}$ such that $||\mathbf{u}||_\lambda < \infty$, where

$$(\mathbf{u}, \mathbf{v}) = \sum_{i \in \mathbb{Z}} \lambda^{-|i|} x_i \cdot y_i,$$

$||\mathbf{u}||_\lambda = \sqrt{(\mathbf{u}, \mathbf{u})}$. Clearly $\mathcal{A}_\lambda$ is metric, complete, separable, and $\tilde{\mathcal{B}}_C$ with the product topology is a subspace of $\mathcal{A}_\lambda$.

Proposition 2.2 *If h satisfies (A1),(E1), then it satisfies (A3).*

Proof: It is easy to check that (E1) implies that the right-hand side of (1) is uniformly Lipschitz on $\mathcal{A}_\gamma$, which (see e.g. [DK74]) implies that (1) generates a continuous flow on $\mathcal{A}_\gamma$. Since ϕ^t is bijection, the semiflow is backward unique. We set $\tilde{\mathcal{A}} = \mathcal{A}_\gamma$; the remaining properties (A3),(i),(ii) and (iii) can be checked easily. $\square$

The conditions (A1),(A2),(E1) and hence (A3) are e.g. satisfied by all standard-like generating functions

$$h(x, y) = (x - y)^2/2 + P(x), \tag{3}$$

where $P : \mathbb{R}^d \to \mathbb{R}$ is any C^2 function such that $P(x + a) = P(x)$ for all $a \in \mathbb{Z}^d$.

Example 2. The twist maps. We now show that the standard setting of the Aubry-Mather theory is sufficient.

(E2) Assume that $d = 1$, and that h satisfies the twist condition, i.e. that

$$h_{12}(x, y) \leq 0, \tag{4}$$

where indices denote partial derivatives.

Since $d = 1$, we can use the partial ordering on $\mathbb{R}^{\mathbb{Z}}$: $\mathbf{u} \leq \mathbf{v}$ if and only if for each $i \in \mathbb{Z}$, $u_i \leq v_i$.

Proposition 2.3 *Assume that h satisfies (A1) and (E2). Then for each $N \in \mathbb{N}$, (1) generates a continuous backward unique semiflow on $\tilde{\mathcal{B}}_N$.*

Proof: Proposition (2.1) implies existence of a local flow ϕ on $\tilde{\mathcal{B}}$.

We now show that if $\mathbf{u}^0 \in \tilde{\mathcal{B}}_N$ and the solution $\mathbf{u}(\tau)$ of (1), (2) exists on $[0, t]$, then $\mathbf{u}(t) \in \tilde{\mathcal{B}}_N$. The twist condition $h_{12} \leq 0$ implies that the solution of (1) is monotone, i.e. that $\mathbf{u}(0) \leq \mathbf{v}(0)$ implies $\mathbf{u}(t) \leq \mathbf{v}(t)$ where defined (see e.g. [Gol92b], Proposition 2). It is easy to check that $\mathbf{u} \in \tilde{\mathcal{B}}_N$ if and only if

$$T_{-N} S\mathbf{u} \leq \mathbf{u} \leq T_N S\mathbf{u}. \tag{5}$$

Assume that $\mathbf{u}^0 \in \tilde{\mathcal{B}}_N$. Now monotonicity of the solution, (5) and the fact that the solution of (1) commutes with T_N, S, imply $T_{-N} S\mathbf{u}(t) \leq \mathbf{u}(t) \leq T_N S\mathbf{u}(t)$, hence $\mathbf{u}(t) \in \tilde{\mathcal{B}}_N$. Since the right-hand side of (1) is uniformly Lipschitz on $l_{\mathbf{u}^0,\infty} \cap \tilde{\mathcal{B}}_N$ in $||.||_{\mathbf{u}^0,\infty}$ norm; we conclude that (1) generates a semiflow ϕ on $\tilde{\mathcal{B}}_N$.

Let $||.||_\gamma$ be the same norm as in Example 1; this norm on $\tilde{\mathcal{B}}_N$ is compatible with the induced product topology. The right-hand side of (1) is uniformly Lipschitz on $\tilde{\mathcal{B}}_N$ in $||.||_\gamma$, which implies that ϕ is a continuous semiflow (see [DK74]).

We now show backward uniqueness. Assume that for some $\mathbf{x}, \mathbf{y} \in \tilde{\mathcal{B}}_N$, $t > 0$, $\phi^t(\mathbf{x}) = \phi^t(\mathbf{y})$. Let t_0 be the smallest such $t \geq 0$. Now $t_0 > 0$ is in contradiction to the local uniqueness of the solution (Proposition 2.1). We deduce that $t_0 = 0$, hence $\mathbf{x} = \mathbf{y}$. $\square$

Now the choice $\tilde{\mathcal{A}} = \tilde{\mathcal{B}}_N$ satisfies all the properties specified as (A3) except $\tilde{\mathcal{B}}_C \subset \tilde{\mathcal{A}}$ for $C > N$. We can, however, set $\tilde{\mathcal{A}} = \tilde{\mathcal{B}}_N$ for N large enough, and prove all the results that will follow.

2.3 Symplectic and Lagrangian maps.

We denote by $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ the d-dimensional torus, and by $\mathbb{A}^d \cong T^*(\mathbb{T}^d) \cong \mathbb{T}^d \times \mathbb{R}^d$ the d-dimensional cylinder. Let (φ, p) be the canonical coordinates on $\mathbb{A}^d$, $v = \sum_{j=1}^d p_j d\varphi_j$ the Liouville 1-form and $w = dv = \sum_{j=1}^d d\varphi_j \wedge dp_j$ the canonical symplectic form on $\mathbb{A}^d$. We denote by $\tilde{\mathbb{A}}^d \cong \mathbb{R}^d \times \mathbb{R}^d$ the cover of $\mathbb{A}^d$ with canonical coordinates (x,p), $\varphi = x \pmod 1$. Given a diffeomorphism F of $\mathbb{A}^d$, We denote by $\tilde{F}$ its lift, $\tilde{F} : \tilde{\mathbb{A}}^d \to \tilde{\mathbb{A}}^d$.

Definition 2.1 *We say that a diffeomorphism F on $\mathbb{A}^d$ is Lagrangian, if it satisfies the following:*

(L1) *F is* **exact symplectic***; i.e. the form $F^*v - v$ is exact;*

(L2) *F is* **strongly monotone***; i.e. if we write $\tilde{F}$ in the form $\tilde{F}(x,p) = (\tilde{F}_1(x,p), \tilde{F}_2(x,p))$, then $r \to \tilde{F}_1(x,r)$ is a diffeomorphism of $\mathbb{R}^d$ for all $x \in \mathbb{R}^d$.*

We say that a diffeomorphism F on $\mathbb{A}^d$ is generated by a C^2 function $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ (or, equivalently, that h is a generating function of F), if its lift $(x,p) \mapsto (x',p')$ satisfies the relations:

$$\begin{aligned} p &= -h_1(x,x') \\ p' &= h_2(x,x'). \end{aligned} \tag{6}$$

We formulate the following additional condition on h:

(P) For all $x \in \mathbb{R}^d$, the functions $y \mapsto h_2(y,x)$ and $y \mapsto h_1(x,y)$ are diffeomorphisms of $\mathbb{R}^d$.

Proposition 2.4 *Each Lagrangian diffeomorphism F on $\mathbb{A}^d$ is generated by a C^2 function h satisfying (A1), (P). Each C^2 function h satisfying (A1), (P), is a generating function of a Lagrangian diffeomorphism F on $\mathbb{A}^d$.*

Proof: [Her89], section 8. $\square$

A Lagrangian map F whose generating function satisfies (A2) is called monotone, globally positive (terminology by Herman). A sufficient condition on a Lagrangian map to be monotone, globally positive is given in [Her89], 8.12. We will need only (A1-2) and (P) when proving properties of constructed invariant measures; these are the same assumptions as those in [Her89].

2.4 Stationary configurations.

Given a semiflow ψ on $\mathcal{X}$, we say that a point $x \in \mathcal{X}$ is stationary, if for each $t > 0$, $\psi^t(x) = x$. The stationary points of the semiflow ϕ generated by (1) are $\mathbf{u} \in \mathcal{A}$ satisfying $\nabla h(\mathbf{u}) = 0$, i.e. for each $i \in \mathbb{Z}$,

$$h_2(u_{i-1}, u_i) + h_1(u_i, u_{i+1}) = 0. \tag{7}$$

We denote by $\mathcal{S}$ the set of all stationary points in $\mathcal{A}$.

If the function h satisfies (P), then (6) implies that for any $C > 0$, the shift S on the set $\mathcal{S} \cap \mathcal{B}_C$ is conjugate to the Lagrangian map F generated by h, restricted to the set of points $(\varphi, p) \in \mathbb{A}^d$ such that

$$\sup_{n \in \mathbb{Z}} (\tilde{F}_1^n(x,p) - \tilde{F}_1^{n-1}(x,p)) \le C, \tag{8}$$

where $\tilde{F} = (\tilde{F}_1, \tilde{F}_2)$ is a lift of F, and $x = \varphi \pmod 1$. Therefore in the following we restrict our attention to the dynamical system shift S on $\mathcal{S} \cap \mathcal{B}_C$ for C large enough.

Remark 2: The orbits of F not satisfying (8) for any $C > 0$ can fill a significant part of the phase portrait (if F is "far" from integrable), for results on such orbits when $d = 1$ see [Sli99a]. The techniques of this paper could be extended to these orbits, adding two additional assumptions: (i) that $\mathcal{A}$ contains all the solutions of (7); and (ii) that the topology on $\mathcal{S}$ induced by $\mathcal{A}$ is the same as the induced product topology on $\mathcal{S}$.

3 The induced semiflow on the space of measures

The main tool in the following will be the induced semiflow on the space of Borel probability measures on $\mathcal{A}$. In this section we define the induced semiflow, show that it is gradient, and that each measure of interest is contained in a compact set invariant for the semiflow.

Let $\mathcal{A}^*$ denote the set of all Borel probability measures on $\mathcal{A}$. We always assume the weak*-topology on $\mathcal{A}$.

It is important to note that we study simultaneously two dynamical systems on $\mathcal{A}$: the shift S, which corresponds to the dynamics of the Lagrangian map; and the semiflow ϕ. We say that a subset $\mathcal{X}$ is S-invariant, (respectively ϕ-invariant), if $S(\mathcal{X}) = S^{-1}(\mathcal{X}) = \mathcal{X}$ (respectively if for each $t \geq 0$, $\phi^{-t}(\mathcal{X}) = \mathcal{X}$). Analogously, we say that a measure $\mu \in \mathcal{A}^*$ is S-invariant, if $\mu \circ S^{-1} = \mu = \mu \circ S$. A measure $\mu \in \mathcal{A}^*$ is S-ergodic, if for each S-invariant Borel-measurable set $\mathcal{X} \subset \mathcal{A}$, $\mu(\mathcal{X}) = 0$ or 1.

The theory of ϕ-invariant or ergodic measures will not be needed in the following; we only note that all ϕ-invariant measures are supported on $\mathcal{S}$, which was in a more general framework of "extended gradient systems" proved in [Sli99b].

Now we list further notation:

$\mathcal{A}^*_S$, $\mathcal{A}^*_E$ - all S-invariant (resp. S-ergodic) measures in $\mathcal{A}^*$,
$\mathcal{S}^*_S$, $\mathcal{S}^*_E$ - all S-invariant (resp. S-ergodic) measures supported on $\mathcal{S}$

(we say that a measure μ is supported on $\mathcal{S}$, if $\mu(\mathcal{S}) = 1$).

The induced semiflow ϕ^* on $\mathcal{A}^*$ is defined as:

$$\phi^{*t}(\mu) = \mu \circ \phi^{-t}.$$

3.1 Properties of the weak* topology.

(A3) implies that $\mathcal{A}^*$ is complete, metrisable and separable ([Par67]), and all the measures in $\mathcal{A}^*$ are regular. Given a continuous function $f : \mathcal{A} \to \mathcal{A}$, the definition of weak*-topology implies immediately that the induced function $f^* : \mathcal{A}^* \to \mathcal{A}^*$, $f^*(\mu) = \mu \circ f^{-1}$ is weak*-continuous. We conclude that ϕ^{*t} is weak*-continuous for each $t \geq 0$.

Given a sequence of Borel measurable functions $f_n : \mathcal{X} \to \mathcal{X}$, where $\mathcal{X}$ is a metrisable space, and a probability measure μ on $\mathcal{X}$, we say that f_n converges in probability, if for each $\epsilon > 0$, $\mu\{x : d(f_n(x), f(x)\} > \epsilon) \to 0$, where d is a metric on $\mathcal{X}$. A sequence f_n converges in distribution, if $\mu \circ f_n$ weak*-converges to $\mu \circ f$. The following is a well-known fact (see e.g. [Bil68]).

Proposition 3.1 *Given a sequence $f_n : \mathcal{X} \to \mathcal{X}$ of Borel-measurable functions on a complete separable metrisable space $\mathcal{X}$, and a Borel probability measure μ on $\mathcal{X}$, if the sequence f_n converges μ-a.e. to f, it converges to f in probability. If the sequence f_n converges in probability to f, it converges to f in distribution.*

Given a sequence of positive times $t_n \to t$, we conclude that $\phi^{*t_n}\mu \to \phi^{*t}\mu$ in weak*-topology; hence $t \mapsto \phi^{*t}\mu$ is continuous. We conclude the following:

Proposition 3.2 *The semiflow ϕ^* is a continuous semiflow on $\mathcal{A}^*$.*

Since S is continuous and $\mathcal{S}$ closed, $\mathcal{A}^*_S$, $\mathcal{S}^*_S$ are weak*-closed and ϕ^*-invariant sets. In the following, we restrict our attention to the semiflow ϕ^* on $\mathcal{A}^*_S$.

3.2 The semiflow ϕ^* is gradient.

Definition 3.1 *A semiflow ψ on a set $\mathcal{A}$ is strictly gradient, if there exists a function $L : \mathcal{A} \to \mathbb{R}$ (called Lyapunov function), such that for each non-stationary $x \in \mathcal{A}$, and any $t > 0$, $A(\psi(t, x)) < A(x)$.*

One of main difficulties when applying the semiflow ϕ to construction of stationary configurations with given properties, is that it is only formally gradient (the natural "Lyapunov" function $A(\mathbf{u}) = \sum h(u_i, u_{i+1})$ is divergent; for discussion of various properties of such systems we again refer the reader to [Sli99b]). We can, however, overcome this difficulty by working in the space $\mathcal{A}^*_S$.

Theorem 3.3 *The semiflow ϕ^* is strictly gradient on $\mathcal{A}^*_S$, with Lyapunov function*

$$A(\mu) = \int_{\mathcal{A}} h(u_0, u_1) d\mu.$$

The set of stationary points of ϕ^* is $\mathcal{S}^*_S$.

Proof: Clearly $\mathcal{S}^*_S$ is a subset of the set of stationary points of ϕ^*. Assume that $\mu \in \mathcal{A}^*_S$ is a measure not supported on $\mathcal{S}$. It is sufficient to show that A is strictly decreasing along the ϕ^*-orbit of μ.

In the following, we use the notation $x_i^t = \phi^t(\mathbf{x})_i$. Given $T > 0$, we calculate, using the S-invariance of μ in the fourth line below:

$$\begin{aligned}
A(\phi^{*T}(\mu)) - A(\mu) &= \int_{\mathcal{A}} h(x_0, x_1) d\phi^{*T}\mu - \int_{\mathcal{A}} h(x_0, x_1) d\mu \\
&= \int_{\mathcal{A}} h(x_0^T, x_1^T) d\mu - \int_{\mathcal{A}} h(x_0, x_1) d\mu \\
&= \int_{\mathcal{A}} \int_0^T (h_1(x_0^t, x_1^t)\dot{x}_0 + h_2(x_0^t, x_1^t)\dot{x}_1) dt d\mu \\
&= \int_{\mathcal{A}} \int_0^T h_1(x_0^t, x_1^t)\dot{x}_0 dt d\mu + \int_{\mathcal{A}} \int_0^T h_2(x_0^t, x_1^t)\dot{x}_1 dt d(\mu \circ S) \\
&= \int_{\mathcal{A}} \int_0^T (h_1(x_0^t, x_1^t)\dot{x}_0 + h_2(x_{-1}^t, x_0^t)\dot{x}_0) dt d\mu \\
&= \int_{\mathcal{A}} \int_0^T -(\dot{x}_0^t)^2 dt d\mu. \qquad (9)
\end{aligned}$$

Assume that the right-hand side above is equal to 0. Then $\int_0^T (\dot{x}_0^t)^2 dt = 0$ μ-a.e., and then $\phi^t(\mathbf{x})_0 = \phi(\mathbf{x})_0$ μ-almost everywhere, for all $t \in [0, T]$. Since μ is S-invariant, $\phi^t(\mathbf{x}) = \phi(\mathbf{x})$ for each $t \in [0, T]$, μ a.e., which is a contradiction to the assumption that $\mu \notin \mathcal{S}^*_S$. □

3.3 The action of a measure.

The function $A(\mu) = \int_{\mathcal{A}} h(u_0, u_1)d\mu$ defined in Theorem 3.3 is called the action of μ. Restriction of A to $\mathcal{S}_S^*$ is the same action defined by Mather ([Mat89]) in the case of twist maps, and the analogue of the action for Lagrangian flows ([Mat91]).

Note that for each $\mu \in \mathcal{A}_S^*$, and each $i \in \mathbb{Z}$, the S-invariance of μ implies that $A(\mu) = \int_{\mathcal{A}} h(u_i, u_{i+1})d\mu$.

We define the action $A(\mathbf{x})$ of a configuration $\mathbf{x}$ as $A(\mathbf{x}) = 1/N \cdot \lim_{i=-N}^{N-1} h(x_i, x_{i+1})$, if the expression is convergent. If a measure $\mu \in \mathcal{A}_S^*$ is S-ergodic, and $A(\mu) < \infty$, then for μ-a.e. $\mathbf{x}$, $A(\mathbf{x}) = A(\mu)$.

Since by (A3),(i), $\mathbf{u} \mapsto (u_0, u_1)$ is continuous, so is $\mathbf{u} \mapsto h(u_0, u_1)$. We conclude:

Lemma 3.4 *$A : \mathcal{A}_S^* \to \mathbb{R}$ is lower-semi continuous.*

Proof: Let $A_C(\mu) = \int_{\mathcal{A}} \min(h(x_0, x_1), C)d\mu$, for some $C \in \mathbb{R}$. Since $\min(h(x_0, x_1), C) : \mathcal{A} \to \mathbb{R}$ is bounded and continuous, A_C is continuous, and $A_C \uparrow A$ as $C \to \infty$, which implies the claim. $\square$

Given a real number M, we denote by $\mathcal{X}_M^*$ the set of all $\mu \in \mathcal{A}_S^*$ such that $A(\mu) \leq M$. Theorem 3.3 implies that $\mathcal{X}_M^*$ is ϕ^*-invariant, and the lower semi-continuity of L implies that it is closed.

Conditions (A1) and the continuity of h imply that there exists a constant $K \geq 0$ such that for each $(x, y) \in \mathbb{R}^2$, $h(x, y) \geq -K$. We conclude that for each $\mu \in \mathcal{A}_S^*$,

$$A(\mu) \geq -K. \tag{10}$$

3.4 Compactness of $\mathcal{X}_M^*$.

If we define $\mathcal{B}_C$, $C \in \mathbb{R}$, to be the set of all $\mathbf{u} \in \mathcal{A}$ such that $\sup_{i\in\mathbb{Z}} |u_{i+1} - u_i| \leq C$, then $\mathcal{B}_C$ is homeomorphic to

$$([-C, C]^d)^{\mathbb{N}} \times \mathbb{T}^d \times ([-C, C]^d)^{\mathbb{N}}.$$

The Tychonoff's theorem implies that $\mathcal{B}_C$ is compact. We will show that each $\mu \in \mathcal{X}_M^*$ is "almost" supported on $\mathcal{B}_C$ for large C.

We say that a family Π of Borel probability measures on a metric space $\mathcal{X}$ is tight, if for each ϵ, there exists a compact set $\mathcal{C} \subset \mathcal{X}$ such that for each $m \in \Pi$, $m(C) \geq 1 - \epsilon$. The following is a well known result that we need (proof is in [Bil68], Theorems 6.1 and 6.2).

Theorem 3.5 *A family of Borel probability measures on a metrisable, complete, separable space $\mathcal{X}$ is relatively compact if and only if it is tight.*

Theorem 3.6 *For each $M \in \mathbb{R}$, the set $\mathcal{X}_M^*$ is tight, hence compact.*

Proof: *Step 1.* The condition (A2) implies existence of constants $C, d > 0$ such that for each $x, y \in \mathbb{R}^d$, $|x - y| > d$ implies

$$h(x, y) \geq C|x - y|. \tag{11}$$

Choose ϵ, $0 < \epsilon < \frac{M+K}{Cd}$, where K is the constant from (10). Given $m \in \mathcal{X}_M^*$, we get for any $i \in \mathbb{Z}$, applying (11) and S-invariance of m:

$$\begin{aligned} M \geq \int_{\mathcal{A}/T} h(x_i, x_{i+1})dm &= \int_{|x_{i+1}-x_i| \geq \frac{M+K}{C\epsilon}} h(x_i, x_{i+1})dm \\ &+ \int_{|x_{i+1}-x_i| \leq \frac{M+K}{C\epsilon}} h(x_i, x_{i+1})dm \geq C\frac{M+K}{C\epsilon} m\left(|x_{i+1} - x_i| \geq \frac{M+K}{C\epsilon}\right) - K, \end{aligned}$$

and therefore

$$m\left(|x_{i+1}-x_i| \geq \frac{M+K}{C\epsilon}\right) \leq \epsilon. \tag{12}$$

Step 2. Define the set

$$\mathcal{Y}_\epsilon = \{\mathbf{x} \in \mathcal{A},\ |x_{k+1}-x_k| \leq \frac{M+K}{C\epsilon}(|k|^2+1),\ k \in \mathbb{Z}\}/T.$$

(A3),(ii) implies that $\mathcal{Y}_\epsilon$ is compact, hence measurable. Then, applying (12), we get

$$\begin{aligned} m(\mathcal{Y}_\epsilon) &= 1 - m(\mathcal{Y}_\epsilon^c) \geq 1 - \sum_{k\in\mathbb{Z}} m\left(|x_{i+1}-x_i| \geq \frac{M+K}{C\epsilon}(|k|^2+1)\right) \\ &\geq 1 - \epsilon \sum_{k\in\mathbb{Z}} \frac{1}{|k|^2+1} \geq 1 - \epsilon a, \end{aligned} \tag{13}$$

where $a = \sum_{k\in\mathbb{Z}} 1/(|k|^2+1)$. (A3),(ii) implies that $\mathcal{Y}_\epsilon$ is compact. Since ϵ was arbitrary, Theorem 3.5 implies now that $\mathcal{X}_M^*$ is relatively compact. Lower semi-continuity of A implies that $\mathcal{X}_M^*$ is closed, hence compact. □

Remark 3: The statement of Theorem 3.6 is still true if we replace (A2) with a weaker assumption that for some $C, d > 0$, $|x-y| > d$ implies (11).

We can summarise the results of this section in the following:

Corollary 3.7 *For each $M \in \mathbb{R}$, the set $\mathcal{X}_M^*$ is compact, invariant for the semiflow ϕ^*.*

4 Construction of invariant measures of Lagrangian maps

We denote by $\mathcal{X}^* = \cup_{M\in\mathbb{R}}\mathcal{X}_M^*$ the set of all measures in $\mathcal{A}_S^*$ with finite action. Then $A : \mathcal{X}^* \to \mathbb{R}$ is well-defined and lower semi-continuous. Recall that ϕ^* is a strictly gradient flow on $\mathcal{X}^*$ with the Lyapunov function A, so we can use the following well-known elementary properties of strict gradient semiflows:

Theorem 4.1 *Given a semiflow ψ on a metrisable space $\mathcal{X}$ with a strict Lyapunov function L, every local minimum of L on $\mathcal{X}$ is a stationary point of ψ.*

If for each $M \in \mathbb{R}$, the set $\{x \in \mathcal{X},\ :\ L(x) \leq M\}$ is compact, then for each $x \in \mathcal{X}$, the ω-limit set $\omega(x)$ is non-empty and consists of stationary points.

4.1 Construction by minimisation.

Theorem 4.2 *Let $\mathcal{Y}^*$ be a closed, non-empty, ϕ^*-invariant subset of of $\mathcal{X}^*$. Then the set of local minima of $A|_{\mathcal{Y}^*}$ is a non-empty subset of $\mathcal{S}_S^*$. In particular, A attains its minimum on $\mathcal{Y}^*$.*

Proof: Since $\mathcal{Y}^*$ is ϕ^*-invariant, and A is a strict Lyapunov function for the semiflow ϕ^*, every local minima of A is a stationary point of ϕ^*. Theorem 3.3 implies that the set of local minima of $A|_{\mathcal{Y}^*}$ is a subset of $\mathcal{S}_S^*$. Since for all $M > 0$, $\mathcal{Y}^* \cap \mathcal{X}_M^*$ is compact and L lower semi-continuous, it attains its minimum on $\mathcal{Y}^*$. □

We construct measures in $\mathcal{S}_S^*$ using the minimisation method in section 6.

4.2 Construction by relaxation.

In the following, $\omega(\mu)$ denotes the ω-limit set of a measure $\mu \in \mathcal{X}^*$ with respect to the semiflow ϕ^*.

Theorem 4.3 *Given any $\mu \in \mathcal{X}^*$, $\omega(\mu)$ is non-empty and $\omega(\mu) \subset \mathcal{S}_S^*$.*

Proof: Theorem 3.6 implies that the set $\{\mu \in \mathcal{X}^* \ : \ A(\mu) \leq M\} = \mathcal{X}_M^*$ is compact; now Theorems 3.3 and 4.1 imply the claim. $\square$

We construct measures in $\mathcal{S}_S^*$ using the relaxation method in section 7.

Remark 4: *Construction by the fixed point argument.* Given a closed, ϕ^*-invariant subset $\mathcal{Y}^*$ of $\mathcal{X}^*$, the third possible method of construction of a measure $\mathcal{S}_S^* \subset \mathcal{Y}^*$ is to apply the Tychonoff fixed-point theorem to the function ϕ^{*1} on the convex hull of $\mathcal{Y}^*$, and an application of the Choquet representation theorem and the Krein-Milman theorem then implies that $\mathcal{S}_S^* \cap \mathcal{Y}^* \neq \emptyset$. Since we do not know an example of a new result obtained with this method, details are omitted, and can be found in [Sli99c].

5 Invariants of the induced semiflow

In this section we show that all ergodic-theoretical properties are ϕ^*-invariant. We discuss as well which properties are preserved in the limit; i.e. whether they extend to the closure of the ϕ^*-orbit of a measure.

5.1 Ergodicity.

We first prove that ergodicity is an invariant of the semiflow ϕ^*. This will follow from a more general Theorem 5.3 below, but we first give an elementary proof.

Lemma 5.1 *(i) for each $\mathcal{U} \subset \mathcal{A}$, $\phi^{-t}\phi^t(\mathcal{U}) = \phi^t\phi^{-t}(\mathcal{U}) = \mathcal{U}$;*
(ii) for each $\mathcal{U} \subset \mathcal{A}$, $\phi^{-t}S(\mathcal{U}) = S\phi^{-t}(\mathcal{U})$;
(iii) if $\mathcal{U}$ is S-invariant, so is $\phi^{-t}(\mathcal{U})$.

Proof: (i) follows from injectivity of ϕ^t, assumed in (A3),(iv). Applying (i), we get:

$$S\phi^{-t}(\mathcal{U}) = \phi^{-t}\phi^t S\phi^{-t}(\mathcal{U}) = \phi^{-t}S\phi^t\phi^{-t}(\mathcal{U}) = \phi^{-t}S(\mathcal{U});$$

(iii) follows from (ii). $\square$

Proposition 5.2 *The set $\mathcal{A}_E^*$ is invariant for semiflow ϕ^*.*

Proof: Choose an S-ergodic measure μ, and an S-invariant set $\mathcal{U}$. Proposition 5.1, (iii) implies that $\phi^{-t}(\mathcal{U})$ is S-invariant, hence $\phi^{*t}\mu(\mathcal{U}) = \mu(\phi^{-t}(\mathcal{U})) \in \{0,1\}$. Therefore $\phi^{*t}\mu$ is S-ergodic. $\square$

The set of S-ergodic measures is not in general weak*-closed, which will be the major difficulty when constructing measures in $\mathcal{A}_E^*$ with prescribed properties.

Remark 5: The following conjecture is still open: if $\mu \in \mathcal{X}^*$ is S-ergodic, then each $m \in \omega(\mu)$ is S-ergodic. If it is true, it would in particular imply immediately existence of quasiperiodic orbits with an arbitrary rotation vector (see section 7 for details). A counterexample to an analogous claim in the case of Lagrangian flows is the measure constructed in [Sli99b], Example 8.

5.2 Ergodic invariants.

Recall that, given probability spaces $(X_1, \mathcal{B}_1, m_1)$ and $(X_2, \mathcal{B}_2, m_2)$, and measure-preserving transformations $S_1 : X_1 \to X_1$, $S_2 : X_2 \to X_2$, we say that the dynamical system $(X_2, \mathcal{B}_2, m_2, S_2)$ is a factor of the dynamical system $(X_1, \mathcal{B}_1, m_1, S_1)$, if there exist measurable sets $M_1 \subset X_1$, $M_2 \subset X_2$ of full measure, S_1- (resp. S_2-) invariant, and a measurable, measure-preserving map $\theta : M_1 \to M_2$ (called factor map), such that

$$\theta \circ S_1 = S_2 \circ \theta.$$

We say that $(X_1, \mathcal{B}_1, m_1, S_1)$ and $(X_2, \mathcal{B}_2, m_2, S_2)$ are isomorphic, if there exists a factor map which is a bijection, with measurable inverse.

Invariant measures of isomorphic dynamical systems have a number of properties (spectral invariants) in common, to be specified in the following. We first prove that the flow ϕ^t is the isomorphism. (Below $\mathcal{B}$ denotes the Borel σ-algebra on $\mathcal{A}$.)

Theorem 5.3 *For each measure $\mu \in \mathcal{A}_S^*$ and $t \geq 0$, the dynamical systems $(\mathcal{A}, \mathcal{B}, \mu, S)$ and $(\mathcal{A}, \mathcal{B}, \phi^{*t}\mu, S)$ are isomorphic.*

Proof: Let $M_1 = \mathcal{A}$, and $M_2 = \phi^t(\mathcal{A})$. The map $\theta := \phi^t|_{M_1}$, $\theta : M_1 \to M_2$ is measure-preserving (because of the definition of $\phi^{*t}\mu$), measurable (since it is continuous), surjective, injective (because of (A3),(iv)) and therefore bijective. Since $\mathcal{A}$ is separable, metrisable and complete, the Kuratovski theorem (see [Par67], Theorem 3.9) implies that the inverse of θ is measurable. Since ϕ^t commutes with S, so does θ. We conclude that θ is isomorphism. $\square$

Standard results in ergodic theory (see [Wal82] for definitions; and Theorems 2.13 and 4.11 therein for proofs) now imply:

Corollary 5.4 *The flow ϕ^* preserves the following properties and invariants of a measure $\mu \in \mathcal{A}_S^*$ with respect to S: (i) the property of weak-mixing; (ii) the property of strong-mixing; (iii) the Kolmogorov-Sinai entropy; (iv) spectrum.*

Remark 6: None of the properties above are necessarily preserved in the limit. An example follows: assume $d = 1$, and $h(x, y) = (x - y)^2$. In this (integrable) case, all measures in $\mathcal{S}_S^*$ have 0 entropy and are not mixing. We can easily find a strongly-mixing measure with positive entropy in $\mathcal{X}^*$, say the Bernoulli measure μ_B supported on the set of configurations $\{\mathbf{x} \ : \ x_i = 0 \text{ or } 1\}$; and then $\omega(\mu_B) \subset \mathcal{S}_S^*$.

5.3 Linearity of the flow.

The definition of the semiflow ϕ^* implies that the map ϕ^{*t} is affine, i.e. that, given $p, q \geq 0$, $p + q = 1$, $\phi^{*t}(p\mu_1 + q\mu_2) = p\phi^{*t}\mu_1 + q\phi^{*t}\mu_2$.

Given an S-invariant measure $\mu \in \mathcal{A}_S^*$, the Choquet representation theorem (see [Phe66]) implies that there exists a measure τ on the Borel subsets of the space $\mathcal{A}_S^*$, such that $\tau(\mathcal{A}_E^*) = 1$, and such that for each $f \in L_1(\mathcal{A}, \mu)$,

$$\int_{\mathcal{A}} f(x) \mathrm{d}\mu(x) = \int_{\mathcal{A}_S^*} \left(\int_{\mathcal{A}} f(x) \mathrm{d}m(x) \right) d\tau(m). \tag{14}$$

The measure τ is called an ergodic decomposition of μ.

The definition of Lebesgue integral and Proposition 5.2 imply the following:

Corollary 5.5 *If τ is an ergodic decomposition of $\mu \in \mathcal{A}_S^*$, then $\tau \circ (\phi^*)^{-t}$ is an ergodic decomposition of $\phi^{*t}\mu$, i.e. for each $f \in L_1(\mathcal{A}, \phi^{*t}\mu)$,*

$$\int_{\mathcal{A}} f(x) d\phi^{*t}\mu(x) = \int_{\mathcal{A}_S^*} \left(\int_{\mathcal{A}} f(x) d\phi^{*t} m(x) \right) d\tau(m). \tag{15}$$

5.4 Rotation vectors.

The rotation vector as defined below is a natural generalisation of the rotation number in the Aubry-Mather theory. Rotation number of a measure was introduced by Mather ([Mat89]).

Definition 5.1 *Given a measure $\mu \in \mathcal{A}_S^*$, we define the rotation vector $\rho(\mu)$ of the measure μ as*

$$\rho(\mu) = \int_{\mathcal{A}} (x_1 - x_0) d\mu.$$

Given $\mathbf{x} \in \mathcal{A}$, we define its rotation vector as

$$\rho(\mathbf{x}) = \lim_{|n| \to \infty} \frac{x_n - x_0}{n},$$

if the expression is convergent.

Birkhoff ergodic theorem implies that, given $\mu \in \mathcal{A}_S^*$ such that $|\rho(\mu)| < \infty$, then μ-almost every $\mathbf{x}$ has well defined rotation vector. If $\mu \in \mathcal{A}_E^*$ and $\rho(\mu) = \rho$, then μ-a.e. configuration $\mathbf{x}$ has rotation vector ρ.

We will need the following result on continuity of ρ (see also [Mat91], Lemma on p177).

Lemma 5.6 *The function $\mu \to \rho(\mu)$ is continuous on $\mathcal{X}_M^*$, for all $M > 0$.*

Proof: Applying (A2), for each $\epsilon > 0$, we can find $C(\epsilon) > 0$ such that $|x_1 - x_0| \geq C(\epsilon)$ implies $|h(x_0, x_1)|/|x_1 - x_0| \geq 1/\epsilon$. We define

$$\rho_\epsilon(\mathbf{x}) = \min\left\{1, \frac{C(\epsilon)}{|x_1 - x_0|}\right\} (x_1 - x_0).$$

Then $\rho_\epsilon : \mathcal{A} \to \mathbb{R}^d$ is bounded and continuous, hence $\rho_\epsilon^* : \mu \mapsto \int \rho_\epsilon d\mu$ is continuous. Given $\mu \in \mathcal{X}_M^*$, applying the bound $\int_{\mathcal{A}} |h(x_0, x_1)| d\mu \leq M + K$ (where K is the bound from (10)), we get:

$$\begin{aligned}
|\rho(\mu) - \rho_\epsilon^*| &\leq \int_{\mathcal{A}} |x_1 - x_0 - \rho_\epsilon| d\mu \leq \int_{|x_1 - x_0| \geq C(\epsilon)} |x_1 - x_0| d\mu \leq \int_{|x_1 - x_0| \geq C(\epsilon)} |x_1 - x_0| d\mu \\
&\leq \int_{|x_1 - x_0| \geq C(\epsilon)} \epsilon |h(x_0, x_1)| d\mu \leq \epsilon(M + K). \qquad (16)
\end{aligned}$$

We conclude that $\rho : \mathcal{X}_M^* \to \mathbb{R}^d$ can be uniformly approximated by continuous functions ρ_ϵ^*, hence it is continuous. $\square$

Note that relation (16) implies that for each $\mu \in \mathcal{X}^*$, $|\rho(\mu)| < \infty$. We now prove that the rotation vector is invariant for the semiflow ϕ^*.

Theorem 5.7 *If $\mu \in \mathcal{X}^*$, then for each $t \geq 0$, $\rho(\phi^{*t}\mu) = \rho(\mu)$.*

Proof: First, assume that $\mu \in \mathcal{A}_E^*$. Note that (10) implies that $A(\mu)$ is bounded from below, $A(\mu) \geq -K$. We use relation (9), and for each $t > 0$, $i \in \mathbb{Z}$, we get

$$\int_{\mathcal{A}} \int_0^t (\dot{x}_i(\tau))^2 \mathrm{d}\tau \mathrm{d}\mu \leq A(\mu) + K.$$

The inequality of arithmetic and geometric mean for integrals implies that for each $t > 0$, $i \in \mathbb{Z}$,

$$\int_{\mathcal{A}} \frac{1}{t}(x_i(t) - x_i(0))^2 \mathrm{d}\mu \leq A(\mu) + K.$$

Given $t > 0$, we conclude that for each $\epsilon > 0$ we can find $C > 0$ large enough such that for each $n \in \mathbb{N}$, $\mu(|x_n(t) - x_n(0)| \leq C) \geq 1 - \frac{\epsilon}{2}$, hence

$$\mu(|x_n(t) - x_n(0)| \leq C, |x_0(t) - x_0(0)| \leq C) \geq 1 - \epsilon,$$

$$\text{and} \quad \mu\left(\left|\frac{x_n(t) - x_0(t)}{n} - \frac{x_n(0) - x_0(0)}{n}\right| \leq \frac{2C}{n}\right) \geq 1 - \epsilon. \tag{17}$$

Since $\phi^{*t}\mu$ is ergodic, $\phi^{*t}\mu$-a.e. configuration $\mathbf{x}$ has the rotation number $\rho' = \rho(\phi^{*t}\mu)$. Since ϵ can be arbitrarily small, and n arbitrarily large, (17) implies $\rho' = \rho(\mu)$.

If $\mu \in \mathcal{A}_S^*$, relation (15) implies the claim. □

Now Lemma 5.6 and Theorem 5.7 imply the following:

Corollary 5.8 *For each $\mu \in \mathcal{X}^*$, and any $m \in \omega(\mu)$, $\rho(\mu) = \rho(m)$.*

6 Minimising measures

In this section we prove analogues of Mather's results on minimising measures of Lagrangian flows from [Mat91]. The novelty is the method of construction of minimising measures: their existence is a direct corollary of the fact that the flow ϕ^* is gradient, and in particular does not depend on (P) ((P) is the analogue of the condition of the positive definiteness of Lagrangian flows, used by Mather). Most of the remaining proofs (except partially the proof of the generalisation of the Birkhoff theorem) follow closely [Mat91]; we present only the proofs which are not completely analogous.

6.1 Existence of minimising measures.

We denote by $\mathcal{X}_\rho^*$ the set of all measures $\mu \in \mathcal{X}^*$ (i.e. with finite action $A(\mu)$) such that $\rho(\mu) = \rho$. We first show that $\mathcal{X}_\rho^*$ is non-empty. The function $\tilde{\pi}_\rho : \mathbb{R}^d \to \tilde{\mathcal{A}}$ defined as

$$\tilde{\pi}_\rho(x)_k = x + k\rho, \quad k \in \mathbb{Z}, \tag{18}$$

commutes with translations $x \mapsto x + a$ and T_a, on $\mathbb{R}^d$ and $\tilde{\mathcal{A}}$ (i.e. $\tilde{\pi}_\rho(x + a) = T_a(\tilde{\pi}_\rho(x))$ for all $a \in \mathbb{Z}^d$, hence we can define $\pi_\rho : \mathbb{T}^d \to \mathcal{A}$ such that its lift is $\tilde{\pi}_\rho$. Since $\pi_\rho(\mathbb{T}^d) \subset \mathcal{B}_{|\rho|}$, (A3) implies that π_ρ is continuous, hence Borel-measurable. If R_ρ is the $x \mapsto x + \rho$ translation on the torus $\mathbb{T}^d$, (18) implies that $\pi_\rho \circ R_\rho = S \circ \pi_\rho$. We now define the measure $\mu_\rho = \lambda \circ \pi_\rho^{-1}$, where λ is the Lebesgue-Haar measure on $\mathbb{T}^d$. Since λ is R_ρ-invariant, we deduce that $\mu_\rho \in \mathcal{A}_S^*$. Since supp $\mu_\rho \subset \mathcal{B}_{|\rho|}$, and since $\mathbf{x} \mapsto h(x_0, x_1)$ is continuous, hence bounded on $\mathcal{B}_{|\rho|}$, $A(\mu_\rho) < +\infty$ and $\mu_\rho \in \mathcal{X}_\rho^*$.

Theorem 6.1 *The function A attains its minimum on $\mathcal{X}_\rho^*$ for any $\rho \in \mathbb{R}^d$, and every such minimum is a measure in $\mathcal{S}_S^*$.*

Proof: Let $M = A(\mu_\rho)$. Corollary 3.7, Theorem 5.7 and Lemma 5.6 imply that the set $\mathcal{Y}^* = \mathcal{X}_M^* \cap \mathcal{X}_\rho^*$ is ϕ^*-invariant and compact, and the choice of M implies that it is non-empty. Theorem 4.2 now implies the claim. $\square$

We denote the set of all measures which minimise A on $\mathcal{X}_\rho^*$ by $\mathcal{M}_\rho^*$, and let $\mathcal{M}^* = \cup_{\rho\in\mathbb{R}^d} M_\rho^*$. The elements of $\mathcal{M}^*$ are called minimising measures. In the rest of this section, we discuss properties of measures in $\mathcal{M}^*$ and their supports.

6.2 Minimising configurations.

Analogously as in Aubry-Mather theory, we say that a segment $(x_i, ...x_j)$, $i < j$, is minimising, if for each segment $(y_i, ...y_j)$, such that $y_i = x_i$, $y_j = x_j$,

$$h(x_i, x_{i+1}, ..., x_j) \leq h(y_i, y_{i+1}, ..., y_j),$$

where $h(x_i, x_{i+1}, ..., x_j) = \sum_{k=i}^{j-1} h(x_k, x_{k+1})$. A configuration $\mathbf{x}$ is minimising if each its subsegment $(x_i, ..., x_j)$ is minimising.

Lemma 6.2 *For every $K > 0$ there exists $M(K) > K$ such that if $(x_i, ..., x_j)$ is a minimising segment, and $|x_j - x_i|/(j-i) \leq K$, then for $i \leq i' \leq j' \leq j$ we have $|x_{j'} - x_{i'}|/|j' - i'| \leq M(K)$.*

The proof follows from (A2) (i.e. the superlinear growth of h); it is completely analogous to [Mat91], Lemma, p182, and is omitted.

We denote by $\mathcal{M}$ the set of all minimising configurations. (A3),(i) implies that for each $i < j$, $\mathbf{x} \mapsto h(x_i, ..., x_j)$ is continuous, hence $\mathcal{M}$ is closed. By the support supp μ of a measure μ we mean the set of all $\mathbf{x}$ such that every neighbourhood of $\mathbf{x}$ has a positive measure. The following proposition is essentially [Mat91], Proposition 3, but we adjust the proof to our assumptions.

In the proof we will use the following, easily verifiable statement: if $\mu \in \mathcal{M}^*$ is not ergodic, then almost every measure with respect to the ergodic decomposition of μ is in $\mathcal{M}^*$.

Proposition 6.3 *For any $\mu \in \mathcal{M}^*$,* supp $\mu \subset \mathcal{M}$.

Proof: Since $\mathcal{M}$ is closed, it is sufficient to prove the claim for almost every measure with respect to the ergodic decomposition of μ. Therefore we assume without loss of generality that μ is ergodic. Assume that the claim is not true, and choose $\mathbf{x} \in$ supp μ, $\mathbf{x} \notin \mathcal{M}$. Let $(x_i, ..., x_j)$ be a non-minimising subsegment of $\mathbf{x}$ such that $h(x_i, ..., x_j) - h(y_i, ..., y_j) = 2\epsilon > 0$ for some segment $(y_i, ...y_j)$, $y_i = x_i$, $y_j = x_j$. Let $\mathcal{U}$ be a small neighbourhood of $(x_i, ...x_j)$ such that $\mathbf{z} \in \mathcal{U}$ implies $|h(z_i, ..., z_j) - h(x_i, ..., x_j)| \leq \epsilon$. Choose a μ-generic configuration $\mathbf{z}$, i.e. such that

$$\lim_{N\to\infty} \frac{1}{2N} \sum_{k=-N}^{N} \mathbf{1}_{\mathcal{U}}(S^{k(j-i)}\mathbf{z}) \quad = \quad \mu(\mathcal{U})$$

(where $\mathbf{1}_{\mathcal{U}}$ is the characteristic function of $\mathcal{U}$), $A(\mathbf{z}) = A(\mu)$, and $\rho(\mathbf{z}) = \rho(\mu)$. Let $\mathbf{z}^N$ be any configuration such that the segment $(z_{-N}^N, ..., z_N^N)$ is minimising, and such that $z_{-N}^N = z_{-N}$, $z_N^N = z_N$. Furthermore, applying Lemma 6.2, for N large enough we can choose $\mathbf{z}^N \in \mathcal{B}_{M|\rho|+1}$. Let m^N be the measure supported uniformly on $\{S^{-N}\mathbf{z}^N, ..., S^N\mathbf{z}^N\}$, and m a weak*-limit point of the sequence m^N (it exists, because $\mathcal{B}_{M|\rho|+1}$ is compact), and then $m \in \mathcal{A}_S^*$. It is easy to see that $\rho(m) = \rho(\mathbf{z}) = \rho(\mu)$, and $A(m) \leq A(\mu) - \epsilon$, which is in contradiction to $\mu \in \mathcal{M}^*$. $\square$

6.3 Mather's α and β functions.

Following Mather, we define $\beta(\rho) = A(\mu)$, where $\rho \in \mathbb{R}^d$, and $\mu \in \mathcal{M}_\rho^*$.

Lemma 6.4 *β is convex, and has superlinear growth, i.e. $\beta(\rho)/|\rho| \to \infty$ as $|\rho| \to \infty$.*

Proof: Choose $p, q \geq 0$, $p+q=1$, $\rho_1, \rho_2 \in \mathbb{R}^d$ and $m_1 \in \mathcal{M}^*_{\rho_1}$, $m_2 \in \mathcal{M}^*_{\rho_2}$. Then $\beta(p\rho_1 + q\rho_2) \leq A(pm_1 + qm_2) = pA(m_1) + qA(m_2) = p\beta(\rho_1) + q\beta(\rho_2)$, hence β is convex. Superlinear growth follows directly from the definition of β and (A2). $\square$

Let $\alpha : \mathbb{R}^d \to \mathbb{R}$ be the conjugate function of β in the sense of convex analysis, i.e.

$$\alpha(c) = -\min_{\rho \in \mathbb{R}^d} \{\beta(c) - (c, \rho)\}, \tag{19}$$

where $c \in \mathbb{R}^d$, and (c, h) is the canonical scalar product in $\mathbb{R}^d$. From a basic result of convex analysis (see [Roc70]) it follows that, since β is everywhere finite and has superlinear growth, the same is true for α, and $\beta(\rho) = -\min_{c \in \mathbb{R}^d}\{\alpha(c) - (c, \rho)\}$. If we define

$$A_c(\mu) = A(\mu) - (x, \rho(\mu)),$$

where $\mu \in \mathcal{X}^*$, the definitions of α and β imply that $\alpha(c) = -\min\{A_c(\mu) \ : \ \mu \in \mathcal{X}^*\}$. We denote by $\mathcal{M}^{c*}$ the set of measures $\mu \in \mathcal{X}^*$ which minimise $A_c(\mu)$, the discussion above implies that for all $c \in \mathbb{R}^d$, $\mathcal{M}^{c*}$ is non-empty, and that $\mathcal{M}^* = \cup_{\rho \in \mathbb{R}^d} \mathcal{M}^*_\rho = \cup_{c \in \mathbb{R}^d} \mathcal{M}^{c*}$. (Geometrically, if $\mu \in \mathcal{M}^*_\rho$, then $\mu \in \mathcal{M}^{c*}$, where c is the slope of the supporting hyperplane of the epigraph of β at $\rho(\mu)$.)

An important question (discussed further in section 7) is for which $\rho \in \mathbb{R}^d$, there exists a stationary ergodic measure μ such that $\rho(\mu) = \rho$. The following result follows from the fact that ergodic measures are extremal points of the set of invariant measures (see [Mat91], p179 for details).

Proposition 6.5 *Given $\rho \in \mathbb{R}^d$, sufficient conditions for existence of an ergodic measure in $\mathcal{M}^*_\rho$ are:*

(i) $(\rho, \beta(\rho))$ is an extremal point of the epigraph of β;

(ii) there exists $c \in \mathbb{R}^d$ such that $\mathcal{M}^{c*} \subset \mathcal{M}^*_\rho$ (and then $\mathcal{M}^{c*} = \mathcal{M}^*_\rho$).

Finally, it is easy to check that $\mathcal{M}^{c*}$ is a convex set, and that its extremal points are ergodic measures. We write supp $\mathcal{M}^{c*} = Cl \cup_{\mu \in \mathcal{M}^{c*}}$ supp μ (where Cl is the closure).

Corollary 6.6 *There exists $M > 0$ such that* supp $\mathcal{M}^{c*} \subset \mathcal{B}_M$. *In particular,* supp $\mathcal{M}^{c*}$ *is compact.*

Proof: Since $\beta : \mathbb{R}^d \to \mathbb{R}$ has superlinear growth, there exists $K > 0$ such that for each $\mu \in \mathcal{M}^{c*}$, $|\rho(\mu)| \leq K$. If m is an ergodic measure in $\mathcal{M}^{c*}$, then for m-a.e. $\mathbf{x}$, $|\rho(\mathbf{x})| \leq K$, and then Lemma 6.2 implies that for m-a.e. $\mathbf{x}$, and any $i \in \mathbb{Z}$, $|x_{i+1} - x_i| \leq M(K)$. But the set of such orbits is dense in $\mathcal{M}^{c*}$, hence supp $\mathcal{M}^{c*} \subset \mathcal{B}_{M(K)}$. $\square$

6.4 The Birkhoff-Mather Theorem.

The following Lemma is a generalisation of Aubry's Crossing Lemma for twist maps. The statement is analogous to Mather's Crossing Lemma, [Mat91], Lemma, p186 (Mather's condition of positive definiteness of the flow is replaced with (P)).

Lemma 6.7 Aubry-Mather Crossing Lemma. *Assume that h satisfies (P). For every $K > 0$, there exist $\eta, \delta, C > 0$ such that, if (x_{-1}, x_0, x_1) and (y_{-1}, y_0, y_1) are stationary segments such that $|x_0 - y_0| < \delta$, $|x_1 - y_1| > C|x_0 - y_0|$ (or $|x_{-1} - y_{-1}| > C|x_0 - y_0|$), and $|x_i - x_{i-1}| \leq K$, $|y_i - y_{i-1}| \leq K$, $i \in \{0, 1\}$, then there exist $a, b \in \mathbb{R}^d$ such that $|x_0 - b| \leq 1$, $|y_0 - a| \leq 1$, and*

$$h(y_{-1}, y_0, y_1) + h(x_{-1}, x_0, x_1) - h(y_{-1}, a, x_1) - h(x_{-1}, b, y_1) > \eta|x_0 - y_0|^2. \tag{20}$$

Proof: Assume that for some $C>0$, $|x_1-y_1|>C|x_0-y_0|$ (the case $|x_{-1}-y_{-1}|>C|x_0-y_0|$ is analogous), and that $|x_0-y_0|<\delta$, for some $\delta, C>0$ to be specified later. Let C_1^{-1} the uniform bounds on the norm of the inverse of functions $x\mapsto h_1(y,x)$, $x\mapsto h_2(x,y)$, restricted to the set $\{x\ :\ |x-y|\le K+1\}$ (it exists because of (P)). Then

$$|h_1(y_0,x_1)-h_1(y_0,y_1)|\ge C_1|x_1-y_1|\ge C_1C|x_0-y_0|. \tag{21}$$

If we denote by g the function $a\mapsto g(a)=h(y_{-1},a,x_1)$, (21) and $h_2(y_{-1},y_0)+h_1(y_0,y_1)=0$ imply $|g'(y_0)|\ge C_1C(x_0-y_0)$. Let C_3 be the bound on $h_{11}(x,y)$, h_{22} on the set $\{(x,y),\ :\ |x-y|\le K+2\}$, and then $|(g''(a)x,x)|\le C_3x^2$ on the set $\{a\ \ |a-y_{-1}|,|x_1-a|\le K+2\}$. Then if δ small enough (i.e. such that $C_1C\delta/2C_3\le 1$), we choose

$$a=y_0-\frac{g'(y_0)}{|g'(y_0)|}\frac{C_1C}{2C_3}|x_0-y_0|,$$

and then $|a-y_0|\le 1$. Then the Taylor's formula, using all the bounds from above, implies:

$$\begin{aligned} h(y_{-1},y_0,x_1)-h(y_{-1},a,x_1) &= g'(y_0)(y_0-a)+(g''(\xi)(y_0-a),(y_0-a)) \\ &\ge \frac{(C_1C)^2|x_0-y_0|^2}{4C_3}. \end{aligned} \tag{22}$$

We set $b=y_0$, and then

$$h(x_{-1},y_0,y_1)-h(x_{-1},b,y_1)=0. \tag{23}$$

Applying again Taylor's formula, we get:

$$\begin{aligned} h(x_0,x_1)-h(y_0,x_1)+h(x_{-1},x_0)-h(x_{-1},y_0) &= h_1(x_0,x_1)(y_0-x_0) \\ &+(h_{11}(\xi,x_1)(y_0-x_0),(y_0-x_0))+h_2(x_{-1},x_0)(y_0-x_0) \\ &+(h_{22}(x_{-1},\xi')(y_0-x_0),(y_0-x_0))\ge -2C_3(x_0-y_0)^2. \end{aligned} \tag{24}$$

Summing (22), (23), and (24) we get

$$h(y_{-1},y_0,y_1)+h(x_{-1},x_0,x_1)-h(y_{-1},a,x_1)-h(x_{-1},b,y_1)\ge(\frac{C^2C_1^2}{4C_3}-2C_3)|x_0-y_0|^2.$$

We choose now $\delta\le\min\{1,2C_3/C_1C\}$, $C>\sqrt{8}C_3/C_1$ and $\eta<C^2C_1^2/4C_3-2C_3$. We proved the claim in the case $|x_0-y_0|>0$.

When $|x_0-y_0|=0$, with the same choice of δ, C, η, we set (in the case $|x_1-y_1|>0$, the case $|x_{-1}-y_{-1}|>0$ is analogous) $a=y_0-\delta_0 g'(y_0)/|g'(y_0)|$, where $\delta_0>0$ is small enough, and $b=y_0$. Then the left-hand side of (22)>0, (23)=0, (24)=0, and their sum (the left-hand side of (20)) is strictly positive. □

Theorem 6.8 The Birkhoff-Mather Theorem. *Assume that h satisfies (P). Then projection π_0 : supp $\mathcal{M}^{c*}\to\mathbb{T}^d$ is injective, and there exists a constant $C>0$ such that, if $x_0,y_0\in\pi_0(\text{supp } \mathcal{M}^{c*})$, and $\mathbf{x}=\pi_0^{-1}(x_0)$, $\mathbf{y}=\pi_0^{-1}(y_0)$, then*

$$\begin{aligned} |x_1-y_1| &\le C|x_0-y_0|, \\ |x_{-1}-y_{-1}| &\le C|x_0-y_0|. \end{aligned}$$

Proof: Let K be the constant from Corollary 6.6 such that supp $\mathcal{M}^{c*}\subset\mathcal{B}_K$, and $\eta,\delta,C>0$ the constants from Lemma 6.7. Assume that there exist $\mathbf{x},\mathbf{y}\in$ supp $\mathcal{M}^{c*}$ such that $|x_0-y_0|<\delta$ and

$|x_1 - y_1| > C|x_0 - y_0|$ (or $|x_{-1} - y_{-1}| > C|x_0 - y_0|$, the proof is analogous). We apply Lemma 6.7, find a, b such that (20) is satisfied, and denote the left-hand side of (20) by $5\epsilon > 0$.

Step I. We can find $\delta_0 > 0$ such that for $\xi, \xi', z_{-1}, z_1, z'_{-1}, z'_1 \in \mathbb{R}^d$, if $|x_0 - \xi| \leq 1$, $|y_0 - \xi'| \leq 1$, and $|x_{-1} - z_{-1}| < \delta_0$, $|y_{-1} - z'_{-1}| < \delta_0$, $|x_1 - z_1| < \delta_0$, $|y_1 - z'_1| < \delta_0$, then

$$\begin{aligned} |h(x_{-1}, \xi, y_1) - h(z_{-1}, \xi, z'_1)| &\leq \epsilon \\ |h(y_{-1}, \xi', x_1) - h(z'_{-1}, \xi', z_1)| &\leq \epsilon. \end{aligned} \tag{25}$$

We for the moment work in the space $\tilde{\mathcal{A}}$, let $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$ be configurations from the equivalence classes $\mathbf{x}$, resp. $\mathbf{y}$. Let $\tilde{\mathcal{U}}_1$, $\tilde{\mathcal{U}}_2$ be neighbourhoods of $\tilde{\mathbf{x}}$, $\tilde{\mathbf{y}}$ such that $\tilde{\mathbf{z}} \in \tilde{\mathcal{U}}_1$, $\tilde{\mathbf{z}}' \in \tilde{\mathcal{U}}_2$, $|\tilde{x}_0 - \xi| \leq 1$, $|\tilde{y}_0 - \xi'| \leq 1$ imply $|\tilde{x}_{-1} - \tilde{z}_{-1}| < \delta_0$, $|\tilde{y}_{-1} - \tilde{z}'_{-1}| < \delta_0$, $|\tilde{x}_1 - \tilde{z}_1| < \delta_0$, $|\tilde{y}_1 - \tilde{z}'_1| < \delta_0$, and

$$\begin{aligned} |h(\tilde{x}_{-1}, \xi, \tilde{x}_1) - h(\tilde{z}_{-1}, \xi, \tilde{z}_1)| &\leq \epsilon \\ |h(\tilde{y}_{-1}, \xi', \tilde{y}_1) - h(\tilde{z}'_{-1}, \xi', \tilde{z}'_1)| &\leq \epsilon. \end{aligned} \tag{26}$$

Now Lemma 6.7, and relations (25) and (26) imply that, if $\tilde{\mathbf{z}} \in \tilde{\mathcal{U}}_1$, $\tilde{\mathbf{z}}' \in \tilde{\mathcal{U}}_2$, then

$$h(\tilde{z}_{-1}, \tilde{z}_0, \tilde{z}_1) + h(\tilde{z}'_{-1}, \tilde{z}'_0, \tilde{z}'_1) - h(\tilde{z}_{-1}, b, \tilde{z}'_1) - h(\tilde{z}'_{-1}, a, \tilde{z}_1) \geq \epsilon. \tag{27}$$

Step II. We set $\mathcal{U}_1 = \tilde{\mathcal{U}}_1/T$, $\mathcal{U}_2 = \tilde{\mathcal{U}}_2/T$. Let $m_1, m_2 \in \mathcal{M}^{c*}$ be S-ergodic measures such that $m_1(U_1) > 0$, $m_2(U_2) > 0$, and $\mathbf{u}$, $\mathbf{v}$ m_1- (resp. m_2-) generic points in the following sense: $\rho(\mathbf{u}) = \rho(m_1)$, $\rho(\mathbf{v}) = \rho(m_2)$, $A(\mathbf{u}) = A(m_1)$, $A(\mathbf{v}) = A(m_2)$, and

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{2N} \sum_{k=-N}^{N} \mathbf{1}_{\mathcal{U}_1}(S^k(\mathbf{u})) &= \mu(\mathcal{U}_1), \\ \lim_{N\to\infty} \frac{1}{2N} \sum_{k=-N}^{N} \mathbf{1}_{\mathcal{U}_2}(S^k(\mathbf{v})) &= \mu(\mathcal{U}_2). \end{aligned}$$

We now apply Mather's "surgery". Let $\tilde{\mathbf{u}}, \tilde{\mathbf{v}} \in \tilde{\mathcal{A}}$ be configurations from the equivalence classes $\mathbf{u}$, resp. $\mathbf{v}$. We can find increasing sequences m_k, n_k, $k \in \mathbb{Z}$, and sequences $a_k, b_k \in \mathbb{Z}^d$, such that $T_{a_k} S^{m_k} \tilde{\mathbf{u}} \in \tilde{\mathcal{U}}_1$, $T_{b_k} S^{n_k} \tilde{\mathbf{v}} \in \tilde{\mathcal{U}}_2$, and $(m_k - m_{-k})/2k \to \mu(\mathcal{U}_1)$, $(n_k - n_{-k})/2k \to \mu(\mathcal{U}_2)$, as $k \to \infty$. We now find sequences $m'_k, n'_k \in \mathbb{Z}$, $a'_k, b'_k \in \mathbb{Z}^d$, such that

$$\begin{aligned} m'_{2k} - m'_{2k-1} &= m_{2k} - m_{2k-1}, \\ m'_{2k+1} - m'_{2k} &= n_{2k+1} - n_{2k}, \\ n'_{2k} - n'_{2k-1} &= n_{2k} - n_{2k-1}, \\ n'_{2k+1} - n'_{2k} &= m_{2k+1} - m_{2k}, \\ a'_{2k} - a'_{2k-1} &= b'_{2k-1} - b'_{2k} = b_{2k} - a_{2k}, \\ a'_{2k+1} - a'_{2k} &= b'_{2k} - b'_{2k+1} = a_{2k+1} - b_{2k}. \end{aligned}$$

We now define the configurations $\tilde{\mathbf{u}}'$, $\tilde{\mathbf{v}}'$, as

$$\begin{aligned} \tilde{u}'_{m'_{2k-1}+i} &= \tilde{u}_{m_{2k-1}+i} + a'_{2k-1}, & i &= 1, ..., m_{2k} - m_{2k-1} - 1, \\ \tilde{u}'_{m'_{2k}+i} &= \tilde{v}_{n_{2k}+i} + a'_{2k}, & i &= 1, ..., n_{2k+1} - n_{2k} - 1, \\ \tilde{v}'_{n'_{2k-1}+i} &= \tilde{v}_{n_{2k-1}+i} + b'_{2k-1}, & i &= 1, ..., n_{2k} - n_{2k-1} - 1, \\ \tilde{v}'_{n'_{2k}+i} &= \tilde{u}_{m_{2k}+i} + b'_{2k}, & i &= 1, ..., m_{2k+1} - m_{2k} - 1, \end{aligned}$$

and let $\tilde{u}'_{m'_k}$, $\tilde{v}'_{n'_k}$ be such that the segments $(\tilde{u}'_{m'_k-1}, \tilde{u}'_{m'_k}, \tilde{u}'_{m'_k+1})$, resp. $(\tilde{v}'_{n'_k-1}, \tilde{v}'_{n'_k}, \tilde{v}'_{n'_k+1})$ are minimising (i.e. we "switch" between $\mathbf{u}$, $\mathbf{v}$, when they come into $\mathcal{U}_1$, resp. $\mathcal{U}_2$). Let $\mathbf{u}'$, $\mathbf{v}' \in \mathcal{A}$ be the equivalence classes of $\tilde{\mathbf{u}}'$, $\tilde{\mathbf{v}}'$. Now the construction and (27) imply that

$$\begin{aligned} \limsup_{N\to\infty} \tfrac{1}{2N}(h(u'_{-N}, ..., u'_N) + h(v'_{-N}, ..., v'_N) &\leq \tfrac{A(\mu_1)+A(\mu_2)}{2} - \epsilon \tfrac{\mu_1(\mathcal{U}_1)+\mu_2(\mathcal{U}_2)}{2}, \\ \lim_{N\to\infty} \tfrac{1}{2N}((u'_N - u'_{-N}) + (v'_N - v'_{-N})) &= \tfrac{\rho(\mu_1)+\rho(\mu_2)}{2}. \end{aligned} \tag{28}$$

in $\mathcal{A} = \tilde{\mathcal{A}}/T$. Let m_N, $N \in \mathbb{N}$, be the measure uniformly distributed on the set

$$\{S^{-N}\mathbf{u}', S^{-N+1}\mathbf{u}', ..., S^{N}\mathbf{u}', S^{-N}\mathbf{v}', S^{-N+1}\mathbf{v}', ..., S^{N}\mathbf{v}'\}.$$

The construction and Lemma 6.2 imply the existence of $K' > 0$ such that $\mathbf{u}', \mathbf{v}' \in \mathcal{B}_{K'}$, which is a subset of $\mathcal{A}$ and compact, hence the sequence m_N has a limit point in $\mathcal{A}_S$. Relation (28) implies that $\rho(m) = (\rho(\mu_1) + \rho(\mu_2))/2)$, and $A(m) < (A(\mu_1) + A(\mu_2))/2$. But $\mathcal{M}^{c*}$ is convex, hence $\beta((\rho(\mu_1) + \rho(\mu_2))/2) = (\beta(\rho(\mu_1)) + \beta(\rho(\mu_2)))/2 = (A(\mu_1) + A(\mu_2))/2$. The definition of β implies that $A(m) \geq \beta((\rho(\mu_1) + \rho(\mu_2))/2)$, which is a contradiction.

We proved the claim only when $|x_0 - y_0| \leq \delta$. But since the domain of π_0^{-1} is a subset of $\mathbb{T}^d$, which is compact, the claim follows. □

Remark 7: *On Herman's Theorem.* Herman [Her89], Theorem 8.14, proved that, given a monotone, globally positive exact symplectomorphism F, every F-invariant Lagrangian torus is a graph of a Lipschitz function. Note that, if we assume that the set $\mathcal{M}_\rho^{c*}$ is non-empty, in the proof of Birkhoff-Mather theorem we used only (A1), (A2) and (P), which are the same assumptions as those used by Herman. Under the same assumptions, Herman ([Her89], section 12) proved that every invariant Lagrangian torus consists of minimising orbits. Proposition 6.9 below now implies that every F-invariant measure supported on the Lagrangian torus is minimising. If such a measure μ is unique, and supp μ is equal to the Lagrangian torus, Theorem 6.8 is a generalisation of Herman, Theorem 8.14.

6.5 Periodic approximations of minimising measures.

Finally we show that, for every $\rho \in \mathbb{R}^d$, we can find a periodic orbit arbitrarily close (in weak*-topology) to the support of $\mathcal{M}_\rho$; Bernstein and Katok [BK87] proved a similar result for Lagrangian tori and symplectic maps close to integrable.

Proposition 6.9 *Assume that $\rho \in \mathbb{R}^d$, and that $(x^n_{i_n}, x^n_{i_n+1}, ..., x^n_{j_n})$ is a sequence of minimising segments, such that $j_n - i_n \to \infty$, and $(x^n_{j_n} - x^n_{i_n})/(j_n - i_n) \to \rho$, as $n \to \infty$. Then $h(x^n_{i_n}, ..., x^n_{j_n})/(j_n - i_n) \to \beta(\rho)$ as $n \to \infty$.*

The proof is completely analogous to [Mat91], Proposition 1, and is omitted.

We say that a configuration $\tilde{\mathbf{x}} \in \tilde{\mathcal{A}}$ is of type (a, q), $a \in \mathbb{Z}^d$, $q \in \mathbb{N}$, if $T_a S^q \tilde{\mathbf{x}} = \tilde{\mathbf{x}}$. A configuration $\mathbf{x} \in \mathcal{A}$ is of type (a, q), if any $\tilde{\mathbf{x}}$ in the equivalence class $\mathbf{x}$ is of type (a, q).

Let $\mathcal{P}^*_{a,q}$ denote the set of all S-invariant measures supported on S-images of a single configuration of type (a, q). Elements of $\mathcal{P}^*_{a,q}$ are called periodic measures of type a, q. Then $\mathcal{P}^*_{a,q} \subset \mathcal{X}^* \cap \mathcal{A}^*_E$. It is easy to show that $\mathcal{P}^*_{a,q}$ is weak*-closed, and ϕ^*-invariant.

Theorem 4.2 now implies that A attains its minimum on $\mathcal{P}^*_{a,q}$. We denote the set of minima of $A|_{\mathcal{P}^*_{a,q}}$ by $\mathcal{M}^*_{a,q}$. Then every measure in $\mathcal{M}^*_{a,q}$ is stationary, and supported on S-images of a single stationary orbit of type (a, q).

Corollary 6.10 *For every $\rho \in \mathbb{R}^d$, in every weak*-neighbourhood of $\mathcal{M}^*_\rho$ there exists a stationary periodic measure.*

Proof: Choose a sequence $(a_n, q_n) \in \mathbb{Z}^d \times \mathbb{Z}$ such that $a_n/q_n \to \rho$, as $n \to \infty$. We choose a sequence of measures $m_n \in \mathcal{M}^*_{a_n,q_n}$. We can uniformly bound $|a_n/q_n|$ by some constant K, and then Lemma 6.2 implies that for all $n \in \mathbb{N}$, $m_n \in \mathcal{B}^*_{M(K)}$, which is compact; hence m_n has a limit point m. Lemma 5.6 (i.e. continuity of the rotation vector) shows that $\rho(m) = \rho$, and Proposition 6.9 implies that $A(m) = \beta(\rho)$. By the definition of $\mathcal{M}^*_\rho$, we get $m \in \mathcal{M}^*_\rho$, which implies the claim. □

7 The distance functions and quasiperiodic orbits

7.1 The distance and length functions.

We define the following functions $l, d : \mathcal{A} \to [0, \infty]$, with

$$
\begin{aligned}
d(\mathbf{x}) &= \sup_{t \to \infty} |x_0^t - x_0|, \\
l(\mathbf{x}) &= \sup_{t \to \infty} \int_0^t |\dot{x}_0^\tau| d\tau.
\end{aligned}
$$

Since the semiflow ϕ and the projection $\pi_0 : \mathbf{x} \to x_0$ are continuous, functions d, l are measurable. Clearly, $d(\mathbf{x}) \leq l(\mathbf{x})$.

Example 8: Assume (E2) (i.e. the case of twist maps, $d = 1$), and let $\mathbf{x}$, $\mathbf{y}$ be two minimising configurations. Assume further that $\mathbf{x}$, $\mathbf{y}$ are not such that one of them is on the boundary of the same Birkhoff region of instability where the other one lies (an exceptional case). We set $\mathbf{z} = (\max(x_i, y_i))_{i \in \mathbb{Z}}$ (or $\mathbf{z} = (\min(x_i, y_i))_{i \in \mathbb{Z}}$). The main result of [Sli99d] is that $d(\mathbf{z}) < \infty$ if and only if $\mathbf{x}$ and $\mathbf{y}$ are in the same Birkhoff region of instability.

We define functions $d^*, l^* : \mathcal{A}_S^* \to \mathbb{R}$, with

$$
\begin{aligned}
d^*(\mu) &= \int_{\mathcal{A}} d(\mathbf{x})\mu, \\
l^*(\mu) &= \int_{\mathcal{A}} l(\mathbf{x})\mu.
\end{aligned}
$$

Proposition 7.1 *Functions d, l, d^*, l^* are lower semi-continuous.*

Proof: For some $N \in \mathbb{N}$, we define

$$
\begin{aligned}
d_N &= \min\{ \sup_{t \in [0,N]} |x_0^t - x_0|, N\} \\
l_N &= \min\{ \sup_{t \in [0,N]} \int_0^t |\dot{x}_0^\tau d\tau|, N\}.
\end{aligned}
$$

It is easy to check that $d_N, l_N : \mathcal{A} \to \mathbb{R}$ are continuous, bounded, and that $d_N \uparrow d$, $l_N \uparrow l$, hence d, l are lower semi-continuous. We define $d_N^*(\mu) = \int_{\mathcal{A}} d_N(\mathbf{x})d\mu$, $l_N^*(\mu) = \int_{\mathcal{A}} l_N(\mathbf{x})d\mu$, and then d_N^*, l_N^* are continuous. The Lebesgue monotone convergence theorem implies that $d_N^* \uparrow d^*$, $l_N^* \uparrow l^*$ which implies the lower semi-continuity of d^*, l^*. $\square$

7.2 Integrability of d^* and l^*.

Now we show some important consequences of integrability of l^* and d^*; the following results will be tools for construction of invariant measures when using the relaxation technique.

We first need the following technical result; we denote by $Cl\mathcal{O}(\mathbf{x})$ the closure of the ϕ-orbit of $\mathbf{x}$.

Lemma 7.2 *If $\mu \in \mathcal{X}^*$, and $\int_{\mathcal{A}} d^*(\mathbf{x})d\mu < \infty$, then for μ-a.e. $\mathbf{x} \in \mathcal{A}$, $Cl\mathcal{O}(\mathbf{x})$ is compact.*

Proof: Let $\mathcal{Y}_C = \{\mathbf{x} \in \mathcal{A} \ : \ \forall k \in \mathbb{Z},\ |u_{k+1} - u_k| \leq C(k^2 + 1)\}$. Choose $\epsilon > 0$, and the same reasoning as in the proof of Theorem 3.6 implies the existence of $C_\epsilon > 0$ such that $\mu(\mathcal{Y}_{C_\epsilon}) \geq 1 - \epsilon$. Since μ is S-invariant, $\int_{\mathcal{A}} d(S^n\mathbf{x})d\mu = \int_{\mathcal{A}} d(\mathbf{x})d\mu$, and similarly we deduce the existence of a constant D_ϵ and the set $\mathcal{Y}'_{D_\epsilon} = \{\mathbf{x} \in \mathcal{A} \;:\; \forall k \in \mathbb{Z},\; d(S^k\mathbf{x}) \le D_\epsilon(k^2+1)\}$ such that $\mu(\mathcal{Y}'_{D_\epsilon}) \ge 1-\epsilon$. Then if $\mathcal{Z}_\epsilon = \mathcal{Y}_{C_\epsilon} \cap \mathcal{Y}'_{D_\epsilon}$, then $\mu(\mathcal{Z}_\epsilon) \ge 1-2\epsilon$, and for all $\mathbf{x} \in \mathcal{Z}_\epsilon$ and $t \ge 0$, $\mathbf{x}^t \in \mathcal{Y}_{C_\epsilon+2D_\epsilon}$, which is by (A3),(ii) compact. Now the set $\mathcal{Z} = \bigcup_{n\in\mathbb{N}} Z_{1/n}$ is the set such that $\mu(\mathcal{Z}) = 1$, and for all $\mathbf{x} \in \mathcal{Z}$, $Cl\mathcal{O}(\mathbf{x})$ is compact. □

Proposition 7.3 *Assume that $\mu \in \mathcal{X}^*$, and that $l^*(\mu) < \infty$. Then $\omega(\mu)$ consists of a single measure m, and the dynamical system $(\mathcal{A}, \mathcal{B}, m, S)$ is a factor of the dynamical system $(\mathcal{A}, \mathcal{B}, \mu, S)$ (where $\mathcal{B}$ is the Borel σ-algebra on $\mathcal{A}$).*

If μ is S-ergodic, so is m.

Proof: Since μ is S-invariant, for all $i \in \mathbb{Z}$, $\int_{\mathcal{A}} |\dot{x}_i^t| < \infty$, and there exists a set $\mathcal{Z}$, $\mu(\mathcal{Z}) = 1$, such that for each $\mathbf{x} \in \mathcal{Z}$, $n \in \mathbb{Z}$, x_n^t is convergent. Lemma 7.2 implies the existence of a set $\mathcal{Z}' \subset \mathcal{Z}$, $\mu(\mathcal{Z}') = 1$, such that for all $\mathbf{x} \in \mathcal{Z}'$, $Cl\mathcal{O}(\mathbf{x})$ is compact. Therefore for all $\mathbf{x} \in \mathcal{Z}'$, $\mathbf{x}^t$ is convergent. We now define the map $\theta : \mathcal{Z}' \to \mathcal{A}$ as: $\theta(\mathbf{x}) = \lim_{t\to\infty} \mathbf{x}^t$ We claim that θ is Borel-measurable. Indeed, if $\mathcal{U}$ is an open subset of $\mathcal{A}$, then $\theta^{-1}(\mathcal{U}) = \bigcup_{n=1}^\infty \bigcap_{k=n}^\infty \phi^{-k}(\mathcal{U})$, hence $\theta^{-1}(\mathcal{U})$ is a measurable set. Since $\phi^t \to \theta$, μ-a.e., Proposition 3.1 implies that $\phi^{*t}\mu \to m$, $m = \mu \circ \theta^{-1}$, and m is the unique measure in $\omega(\mu)$. It follows immediately that θ is the factor map. A factor of an ergodic dynamical system is always ergodic, hence if μ is S-ergodic, so is m. □

Proposition 7.4 *Assume that $\mu \in \mathcal{X}^*$ is S-ergodic, and that $d^*(\mu) < \infty$. Then for each $m \in \omega(\mu)$, and for m-a.e. $\mathbf{x} \in \mathcal{A}$, $\rho(\mathbf{x}) = \rho(\mu)$.*

If ν is a Borel measure on $\mathcal{S}_S^*$, an ergodic decomposition of $m \in \omega(\mu)$, then for ν-almost every measure m', $\rho(m) = \rho(\mu)$.

Proof: Let $\rho = \rho(\mu)$, and choose $\delta > 0$. We define the set

$$\mathcal{Z}_{n,\delta} = \{\mathbf{x} \in \mathcal{A} \;:\; \left|\frac{x_n - x_0}{n} - \rho\right| \le \delta\},$$

and then it is a closed set, and Birkhoff ergodic theorem implies that for all $\epsilon > 0$,

$$\exists n_0' \text{ such that } \forall n \ge n_0',\; \mu(\mathcal{Z}_{n,\delta}) \ge 1-\epsilon. \tag{29}$$

Let $d^*(\mu) = C$, and then, since μ is S-invariant, for all $k \in \mathbb{Z}$, $\int_{\mathcal{A}} d(S^k(\mathbf{x}))\mathrm{d}\mu = C$, and $\mu(d(S^k(\mathbf{x}) \ge D) \le C/D$. Therefore for each $\epsilon > 0$

$$\exists n_0'' \text{ such that } \forall n \ge n_0'',\; \forall k \in \mathbb{Z},\; \mu\left(\left|\frac{d(S^k(\mathbf{x})) - d(\mathbf{x})}{n}\right| \le \delta\right) \ge 1-\epsilon. \tag{30}$$

Now (29) and (30) imply that for all $n \ge n_0 = \max(n_0', n_0'')$, for all $t \ge 0$, $\phi^{*t}\mu(\mathcal{Z}_{n,3\delta}) \ge 1-3\epsilon$. Since $Z_{n,3\delta}$ is closed, we conclude that

$$\forall\, \epsilon > 0,\; \exists n_0, \text{ such that } \forall n > n_0,\; m(\mathcal{Z}_{n,3\delta}) \ge 1-3\epsilon. \tag{31}$$

By the Birkhoff Ergodic Theorem, m-a.e. $\mathbf{x}$ has well defined rotation vector, and (31) implies that for m-a.e. $\mathbf{x}$, $|\rho(\mathbf{x}) - \rho| \le 3\delta$, and the set $\mathcal{Z}_\delta = \bigcup_{k=1}^\infty \bigcap_{n=k}^\infty \mathcal{Z}_{n,4\delta}$ has full m-measure. The definition of ergodic decomposition implies that for ν-a.e. measure m', $m'(\mathcal{Z}_\delta) = 1$, and then $|\rho(m') - \rho| \le 4\delta$. Since δ can be arbitrarily small, the claim is proved. □

7.3 Quasiperiodic orbits.

The construction in section 6 does not ensure existence of S-ergodic stationary measures with every rotation number $\rho \in \mathbb{R}^d$, and in particular existence of stationary configurations with every rotation number. If $d \geq 2$, the Hedlund counter-example [Hed32] suggests that an attempt to represent every rotation vector with a minimising S-ergodic measure can not in general succeed. We construct here, under some assumptions on the function d, S-ergodic measures with an arbitrary rotation vector $\rho \in \mathbb{R}^d$ (which may be only locally minimal in $\mathcal{X}^*$), and, as a consequence, quasiperiodic orbits with that rotation vector.

We say that a configuration $\mathbf{x} \in \mathcal{A}$ is double-recurrent, if there exist two sequences $m_i \to -\infty$, and $n_i \to \infty$, such that $\lim_{i\to\infty} S^{m_i}\mathbf{x} = \mathbf{x}$, $\lim_{i\to\infty} S^{n_i}\mathbf{x} = \mathbf{x}$.

Theorem 7.5 *Given $\rho \in \mathbb{R}^d$, the following two conditions are equivalent:*
(i) There exists a stationary S-ergodic measure in $\mathcal{X}^*_\rho$;
(ii) there exists a S-ergodic $\mu \in \mathcal{X}^*_\rho$ such that $d^*(\mu) < \infty$.

If the equivalent conditions above are true, then there exists a double-recurrent stationary configuration $\mathbf{x}$ such that $\rho(\mathbf{x}) = \rho$.

Proof: The implication (i)$\Rightarrow$(ii) is trivial, since for every $\mu \in \mathcal{S}^*_S$, $d^*(\mu) = 0$. Assume (ii); Theorem 4.3 implies that we can choose a measure $m \in \omega(\mu)$ supported on $\mathcal{S}$. Then by Proposition 7.4, there exists a measure m', such that $\rho(m') = \rho(\mu)$, and since $m \in \mathcal{S}^*_S$, we can choose $m' \in \mathcal{S}^*_S$. Existence of a double recurrent stationary configuration with rotation vector ρ follows now from Birkhoff ergodic theorem and Poincaré recurrence theorem applied to the stationary S-ergodic measure μ with the rotation vector ρ. □

For example, Theorem 7.5 implies that if for a given $\rho \in \mathbb{R}^d$, $d^*(\mu_\rho) < \infty$ (where μ_ρ is the measure constructed in subsection 6.1), then there exists a quasiperiodic double-recurrent orbit with rotation vector ρ.

In the following theorem, we implicitly assume that there does not exist an S-ergodic minimising measure with rotation number ρ (i.e. that $(\rho, \beta(\rho))$ is not an extremal point of the epigraph of β), otherwise the claim is trivial.

Theorem 7.6 *Given $\rho \in \mathbb{R}^d$, assume that there exists a neighbourhood $\mathcal{U}$ of $\mathcal{M}^*_\rho$ such that $d^*|_{\mathcal{U}} < \infty$. Then for every $\epsilon > 0$, there exists a measure $\mu \in \mathcal{S}^*_E$ such that $\rho(\mu) = \rho$ and $A(\mu) \leq \beta(\mu) + \epsilon$.*

We first construct a measure which satisfies all the properties demanded above except stationarity.

Lemma 7.7 *For every $\epsilon > 0$ and $\rho \in \mathbb{R}^d$, there exists an S-ergodic measure $\mu \in \mathcal{X}^*$ such that $\rho(\mu) = \rho$, and $A(\mu) \leq \beta(\mu) + \epsilon$.*

Proof: *Step I. Construction of a configuration.* We find n_0 large enough such that for $n \geq n_0$, $|h(x, y)| \leq n\epsilon/4$ when $|x - y| \leq M$, and $M = M(|\rho| + 2)$ is the constant from Lemma 6.2. Let $(y^n_0, y^n_1, ..., y^n_n)$, $n \in \mathbb{N}$ be a sequence of minimising segments, such that $y^n_0 = 0$ and $x^n_n = n\rho$. Proposition 6.9 implies that we can find $n > n_0$ and $\mathbf{y} = (y_0, ..., y_n) = (y^n_0, ..., y^n_n)$ such that

$$h(y_0, y_1, ..., y_n)/n \leq \beta(\rho) + \epsilon/2. \tag{32}$$

Now we define a configuration $\mathbf{x}$ in the following way: we set

$$x_{nk+r} = y_r + a_k, \quad k \in \mathbb{Z},\ r \in \{0, 1, ..., n-1\},$$

where we choose $a_k \in \mathbb{Z}^d$ such that

$$|x_{nk} + a_k - nk\rho| < 1.$$

The construction of $\mathbf{x}$ now implies that there exists a constant C such that $\sup |x_i - i\rho| \leq C/2$, which implies that, for all $m \in \mathbb{Z}$, $k \in \mathbb{N}$,

$$\left| \frac{x_{m+k} - x_m}{k} - \rho \right| \leq C/|k|. \tag{33}$$

Since the "cost of switching" between two translates of the segment $\mathbf{y}$ in $\mathbf{x}$ are at most $|h(y_{r-1} + a_{k-1}, y_r + a_{k-1}) - h(y_{r-1} + a_{k-1}, y_0 + a_k)| \leq n\epsilon/2$, relation (32) implies that

$$\limsup_{m\to\infty} h(x_{-m}, x_{-m+1}, ..., x_m)/2m \leq \beta(\rho) + \epsilon. \tag{34}$$

Step II. Construction of the measure. Let μ_m be the uniform measure on the set

$$\{S^{-m}\mathbf{x}, S^{-m+1}\mathbf{x}, ..., S^m\mathbf{x}\}.$$

Since for all m, μ_m is supported on the compact set $\mathcal{B}_M$ (where $M = M(|\rho| + 2)$ is the constant chosen at the beginning), we can find a limit point μ of $(\mu_m)_{m\in\mathbb{Z}}$, which is in $\mathcal{A}_S$. Relation (34) implies that

$$A(\mu) \leq \beta(\rho) + \epsilon. \tag{35}$$

We define the set $\mathcal{Z}_k = \{\mathbf{x} \in \mathcal{A} \ : \ |(x_k - x_0)/k - \rho| \leq C/k\}$, and (33) implies that $\mu_m(\mathcal{Z}_k) = 1$ for all $m \in \mathbb{R}$. Since $\mathcal{Z}_k$ is closed, it follows that for all k, $\mu(\mathcal{Z}_k) = 1$. Let $\mathcal{Z} = \cap_{k\in\mathbb{N}}\mathcal{Z}_k$, and then $\mu(\mathcal{Z}) = 1$. Now (35) and the ergodic decomposition theorem applied to μ imply that we can find an S-ergodic measure μ' such that $A(\mu') \leq \beta + \epsilon$ and $\mu'(\mathcal{Z}) = 1$. The later and Birkhoff ergodic theorem imply that $\rho(\mu') = \rho$. □

Proof of Theorem 7.6: The assumptions imply that for ϵ small enough, we can find a measure m described in Lemma 7.7 such that $d(m) < \infty$. We choose $\mu_1 \in \omega(m)$, and then by Theorem 3.3, $A(\mu_1) \leq \beta(\rho) + \epsilon$, and $\mu_1 \in \mathcal{S}_S^*$. Proposition 7.4 and ergodic decomposition theorem applied to μ_1 now imply that we can find $\mu \in \mathcal{S}_E^*$, such that $\rho(\mu) = \rho$, and $A(\mu) \leq \beta(\rho) + \epsilon$. □

Example 9: *On Hedlund's counter-example and Levi's construction.* We now comment on Levi's construction of orbits with an arbitrary rotation vector in Hedlund's counterexample [Lev97], ignoring for the moment that in this case we deal with a Lagrangian (geodesic in particular) flow, and not a map. Levi constructed a closed family of continuous curves (an analogue to a set of configurations in $\mathcal{A}$ here), called "pseudogeodesics", and showed ([Lev97], Lemma D) that (using the terminology developed here) the function d is uniformly bounded and very small on that set. Furthermore, the set of Levi's pseudogeodesics is rich enough, so that one can construct an S-ergodic measure with an arbitrary rotation vector, supported on that set, and recover the analogue of Theorem 7.5, (ii). One could now slightly improve Levi's construction, and construct invariant measures described in Theorems 7.5 and 7.6, as well as double-recurrent quasi-periodic orbits with an arbitrary rotation vector.

Remark 10: *On the assumptions of Theorem 7.6.* We conjecture that the assumptions of Theorem 7.6 are satisfied in, if not all, then in all but some degenerate cases. Assume the contrary, i.e. that for all measures μ constructed in Lemma 7.7, $d^*(\mu) = \infty$. It would mean that, even though the measure μ has arbitrary small ($\leq \epsilon$) relaxation energy $A(\mu) - A(\omega(\mu))$ available, almost every configuration in supp μ gets arbitrarily far away during the relaxation. Furthermore, a result from [Sli99b] implies that for μ-a.e. $\mathbf{x}$, $\omega(\mathbf{x})$ (with respect to ϕ) is a subset of $\mathcal{S}$; it now implies a degeneracy of the phase space of F in a neighbourhood of supp $\mathcal{M}_\rho^*$ in some sense, yet to be understood.

8 Discussion

8.1 The twist maps and Aubry-Mather theory.

In the following we discuss the case of twist maps on $\mathbb{T} \times \mathbb{R}$, i.e. we assume (E2). We first comment on Golé's alternative proof [Gol92b] of the following fundamental result of Aubry-Mather theory.

Theorem 8.1 The Aubry-Mather theorem. *Assume (E2). Then for every $\rho \in \mathbb{R}$, we can find a configuration $\mathbf{x} \in \mathcal{S}$, such that $\rho(\mathbf{x}) = \rho$.*

An important step in Golé's proof (which was based on the study of the semiflow induced by (1)) is the following claim:

(*) Assume (E2), and $N > 0$. The ω-limit set with respect to the semiflow ϕ of each $\mathbf{x} \in \mathcal{B}_N$ contains a configuration in $\mathcal{S}$.

However, Golé's proof of (*) is incorrect (the statement of [Gol92b], Lemma 2, does not extend to a "neighbourhood of x"). Golé later replaced (*) with a weaker statement, and recovered the proof (Erratum to [Gol92b] will appear in [Gol99]). One can prove (*) using the detailed understanding of the energy flow of the equation (1), the proof is in [Sli99b] ((*) is true even without the assumption (E2), i.e. $d = 1$, if the ϕ-orbit of $\mathbf{x}$ is contained in $\mathcal{B}_C$ for some $C > 0$). We now give a short proof of 8.1 which uses only Proposition 2.3 and Theorem 3.3, and complete the argument from [Gol92b].

Proof of Theorem 8.1: For a given $\rho \in \mathbb{R}$, we define the set $\tilde{\mathcal{Y}}_\rho$, of all $\mathbf{x} \in \mathbb{R}^{\mathbb{Z}}$ such that, given $p \in \mathbb{Z}$ and $q \in \mathbb{N}$, $T_p S^q \mathbf{x} \geq \mathbf{x}$ if $p/q > \rho$, and $T_p S^q \mathbf{x} \leq \mathbf{x}$ if $p/q < \rho$, and let $\mathcal{Y}_\rho = \tilde{\mathcal{Y}}_\rho / T$. Then for each $\mathbf{x} \in \mathcal{Y}_\rho$, $\rho(\mathbf{x}) = \rho$. If N is the smallest integer greater than $|\rho|$, $\mathcal{Y}_\rho$ is a closed subset of $\mathcal{B}_N$, hence compact, and the semiflow ϕ is by Proposition 2.3 well-defined on $\mathcal{Y}_\rho$. Monotonicity of the dynamics (1) (see the proof of Proposition 2.3) implies that $\mathcal{Y}_\rho$ is ϕ-invariant, hence the set $\mathcal{Y}^*_\rho$ of S-invariant measures supported on $\mathcal{Y}_\rho$ is non-empty, compact (because of the compactness of $\mathcal{Y}_\rho$), and ϕ^*-invariant. We choose $\mu \in \mathcal{Y}^*_\rho$, and $m \in \omega(\mu)$, then because of Theorem 3.3 and closedness of $\mathcal{Y}^*_\rho$, $m \in \mathcal{Y}^*_\rho \cap \mathcal{S}^*_S$. Then m-a.e. configuration $\mathbf{x}$ is stationary and in $\mathcal{Y}_\rho$, hence $\rho(\mathbf{x}) = \rho$. $\square$

One can, however, use the remaining results of the paper and recover stronger results of the Aubry-Mather theory including the detailed description of the set of minimising measures and their supports, by proving that the function $\beta : \mathbb{R} \to \mathbb{R}$ is strictly convex. This was already done by Mather, [Mat91], section 6.

We conjecture that the Mather's minimising measure (whose support is the Aubry-Mather's set) is the unique measure in the ω-limit set (with respect to ϕ^*) of the measure μ_ρ, constructed in subsection 6.1. Golé's ghost tori [Gol92a] would then be the attractor of supp μ_ρ (i.e. the set $\bigcap_{T \geq 0} Cl \bigcup_{t \geq T} \phi^t(\text{supp } \mu_\rho)$); this idea might lead to a generalisation of the results from [Gol92a] to more degrees of freedom.

8.2 Open problems.

The results of this paper are a contribution towards the description of the phase-space of the semiflow ϕ^*, as a tool for understanding the phase space of the corresponding Lagrangian system. Here we list further possible directions of this program.

Hyperbolicity. A fascinating relationship between a Lagrangian map F and its "dual" semiflows ϕ, ϕ^* is the result by Aubry, Baesens, MacKay [AMB92] (assuming only (A1) and (E1)): a closure of an orbit of F is uniformly hyperbolic, if and only if the corresponding equilibrium of the flow ϕ is hyperbolic (in $||.||_p$ norm, for any $p \in [1, \infty]$). It would be interesting to find a characterisation of hyperbolic invariant measures of F in terms of the semiflow ϕ^*.

Heteroclinic orbits. We propose the following generalisation of Birkhoff region of instability : it is the set of all configurations in $\mathcal{S}$, which is a subset of a connected component of $\{\mathbf{x} \in \mathcal{A} : d(\mathbf{x}) < \infty\}$ (see Example 8 for the case $d = 1$). The author has already succeeded to prove the existence of a heteroclinic orbit joining any two supp $\mathcal{M}^*_\rho$, supp $\mathcal{M}^*_{\rho'}$ in the same Birkhoff region of instability. Therefore the problem of existence of heteroclinic orbits (and "Arnold's diffusion") can be reduced to the study of the distance function d on $\mathcal{A}$.

The distance function and the Peierls Barrier. Mather ([Mat91]) and Fathi ([Fat99]) constructed heteroclinic orbits in Lagrangian systems, assuming certain topological properties of the Peierls set (i.e. the set where the Peierls barrier is 0). Since the Peierls barrier acts as a "barrier" for the flow ϕ, we conjecture that finiteness of the functions d, d^* is closely related (and possibly equivalent) to these properties.

Measures with positive entropy. We conjecture that each set $\mathcal{M}^*_\rho$ not supported on a Lagrangian torus (+ possibly additional assumptions on the Peierls set, as above) can be weak*-approximated by a positive-entropy invariant measure of the Lagrangian map (generalisation of [For96]); we already have initial results in that direction.

The Morse-Floer homology. One might try to construct the analogue of the Morse-Floer homology for the flow ϕ^*, which might in particular lead to better topological understanding of break-ups of invariant tori, and KAM theory.

Acknowledgments

A part of this work was completed while I was a PhD student in DAMTP, University of Cambridge; and is in [Sli99c]. I would like to thank my supervisor, Robert MacKay for his continuing support. My thanks are due to Dimitris Gatzouras and Anthony Quas for their help with ergodic-theoretical arguments. I also thank the organisers of the Winter School on Smooth Ergodic Theory (participation supported by ESF/PRODYN) and CIME Summer School on Hamiltonian Dynamics (participation supported by UNESCO-ROSTE) for inviting me to these most useful workshops.

References

[AMB92] S. Aubry, R. S. MacKay, and C. Baesens. Equivalence of uniform hyperbolicity for symplectic twist maps and phonon gap for Frenkel-Kontorova models. *Physica D*, 56:123–134, 1992.

[Bil68] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.

[BK87] D. Bernstein and A. Katok. Birkhoff periodic orbits for small perturbations of completely integrable systems with convex hamiltonians. *Invent. Math.*, 88:225–241, 1987.

[DK74] J. L. Dalecki and M. G. Krein. *Stability of Solutions of Differential Equations in Banach Space.* American Mathematical Society, Providence, RI, 1974.

[Fat99] A. Fathi. Orbites hétéroclines et ensemble de Peierls. preprint, 1999.

[For96] G. Forni. Construction of invariant measures supported within the gaps of Aubry-Mather sets. *Ergodic Theory Dynam. Systems*, 16, 1996.

[Gol92a] C. Golé. Ghost circles for twist maps. *Journal of Differential Equations*, 97:140–173, 1992.

[Gol92b] C. Golé. A new proof of Aubry-Mather's theorem. *Math. Z.*, 210:441–448, 1992.

[Gol99] C. Golé. *Symplectic twist maps.* to appear with World Scientific Publishers, 1999.

[Hed32] G. A. Hedlund. Geodesics on a two-dimensional Riemannian manifold with periodic coefficients. *Ann. Math.*, 33:719–739, 1932.

[Her89] M. R. Herman. Inegalités a priori pour des tores lagrangiens invariants par des difféomrphisms symplectiques. *Inst. Hautes Etudes Sci. Publ. Math.*, 70, 1989.

[Lev97] M. Levi. Shadowing property of geodesics in Hedlund's metric. *Ergod. Th. Dynam. Sys*, 17:187–204, 1997.

[Mat89] J. N. Mather. Minimal measures. *Comment. Math. Helv.*, 64(3):375–394, 1989.

[Mat91] J. N. Mather. Action minimising invariant measures for positive definite Lagrangian systems. *Math. Zeit.*, 207:169–207, 1991.

[Par67] K. R. Parthasarathy. *Probability Measures on Metric Spaces.* Academic Press, New York, 1967.

[Phe66] R. Phelps. *Lectures on Choquet's Theorem.* Van Nostrand, N.J., 1966.

[Roc70] R.T. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[Sli99a] S. Slijepčević. Accelerating orbits of twist diffeomorphisms on a torus. *Glasnik Matematički*, 34(1), 1999.

[Sli99b] S. Slijepčević. Extended gradient systems: dimension one. *Disc. Cont. Dynamical Systems*, 1999. to appear.

[Sli99c] S. Slijepčević. Gradient dynamics of Frenkel-Kontorova models and twist maps. PhD thesis, University of Cambridge, 1999.

[Sli99d] S. Slijepčević. Monotone gradient dynamics and Mather's shadowing. *Nonlinearity*, 12:969–986, 1999.

[Wal82] P. Walters. *An Introduction to Ergodic Theory.* GTM 79. Springer, 1982.